SEQUENCE LISTING

<110>   Renz, Andreas
        Heinz, Ernst
        Abbadi, Amine
        Domergue, Frederic
        Zank, Thorsten

<120>   METHOD FOR THE PRODUCTION OF POLYUNSATURATED FATTY ACIDS

<130>   12810-00119-US

<150>   PCT/EP2004/000771
<151>   2004-01-29

<150>   DE 103 08 836.9
<151>   2003-02-27

<160>   57

<170>   PatentIn version 3.3

<210>   1
<211>   849
<212>   DNA
<213>   Caenorhabditis elegans

<220>
<221>   CDS
<222>   (1)..(849)
<223>   Acyl-CoA:lysophospholipid acyltransferase

<400>   1
atg gag aac ttc tgg tcg atc gtc gtg ttt ttt cta ctc tca att ctc   48
Met Glu Asn Phe Trp Ser Ile Val Val Phe Phe Leu Leu Ser Ile Leu
1   5   10   15

ttc att tta tat aac ata tcg aca gta tgc cac tac tat atg cgg att   96
Phe Ile Leu Tyr Asn Ile Ser Thr Val Cys His Tyr Tyr Met Arg Ile
                20   25   30

tcg ttt tat tac ttc aca att tta ttg cat gga atg gaa gtt tgt gtt   144
Ser Phe Tyr Tyr Phe Thr Ile Leu Leu His Gly Met Glu Val Cys Val
            35   40   45

aca atg atc cct tct tgg cta aat ggg aag ggt gct gat tac gtg ttt   192
Thr Met Ile Pro Ser Trp Leu Asn Gly Lys Gly Ala Asp Tyr Val Phe
        50   55   60

cac tcg ttt ttc tat tgg tgt aaa tgg act ggt gtt cat aca aca gtc   240
His Ser Phe Phe Tyr Trp Cys Lys Trp Thr Gly Val His Thr Thr Val
65   70   75   80

tat gga tat gaa aaa aca caa gtt gaa ggt ccg gct gta gtt att tgt   288
Tyr Gly Tyr Glu Lys Thr Gln Val Glu Gly Pro Ala Val Val Ile Cys
                    85   90   95

```
aat cat cag agt tct ctc gac att cta tcg atg gca tca atc tgg ccg        336
Asn His Gln Ser Ser Leu Asp Ile Leu Ser Met Ala Ser Ile Trp Pro
        100             105             110

aag aat tgt gtt gta atg atg aaa cga att ctt gcc tat gtt cca ttc        384
Lys Asn Cys Val Val Met Met Lys Arg Ile Leu Ala Tyr Val Pro Phe
        115             120             125

ttc aat ctc gga gcc tac ttt tcc aac aca atc ttc atc gat cga tat        432
Phe Asn Leu Gly Ala Tyr Phe Ser Asn Thr Ile Phe Ile Asp Arg Tyr
    130             135             140

aac cgt gaa cgt gcg atg gct tca gtt gat tat tgt gca tct gaa atg        480
Asn Arg Glu Arg Ala Met Ala Ser Val Asp Tyr Cys Ala Ser Glu Met
145             150             155             160

aag aac aga aat ctt aaa ctt tgg gta ttt ccg gaa gga aca aga aat        528
Lys Asn Arg Asn Leu Lys Leu Trp Val Phe Pro Glu Gly Thr Arg Asn
            165             170             175

cgt gaa gga ggg ttc att cca ttc aag aaa gga gca ttc aat att gca        576
Arg Glu Gly Gly Phe Ile Pro Phe Lys Lys Gly Ala Phe Asn Ile Ala
            180             185             190

gtt cgt gcg cag att ccc att att cca gtt gta ttc tca gac tat cgg        624
Val Arg Ala Gln Ile Pro Ile Ile Pro Val Val Phe Ser Asp Tyr Arg
        195             200             205

gat ttc tac tca aag cca ggc cga tat ttc aag aat gat gga gaa gtt        672
Asp Phe Tyr Ser Lys Pro Gly Arg Tyr Phe Lys Asn Asp Gly Glu Val
        210             215             220

gtt att cga gtt ctg gat gcg att cca aca aaa ggg ctc act ctt gat        720
Val Ile Arg Val Leu Asp Ala Ile Pro Thr Lys Gly Leu Thr Leu Asp
225             230             235             240

gac gtc agc gag ttg tct gat atg tgt cgg gac gtt atg ttg gca gcc        768
Asp Val Ser Glu Leu Ser Asp Met Cys Arg Asp Val Met Leu Ala Ala
            245             250             255

tat aag gaa gtt act cta gaa gct cag caa cga aat gcg aca cgg cgt        816
Tyr Lys Glu Val Thr Leu Glu Ala Gln Gln Arg Asn Ala Thr Arg Arg
            260             265             270

gga gaa aca aaa gac ggg aag aaa tct gag taa                            849
Gly Glu Thr Lys Asp Gly Lys Lys Ser Glu
            275             280
```

<210>   2
<211>   282
<212>   PRT
<213>   Caenorhabditis elegans

<400>   2

Met Glu Asn Phe Trp Ser Ile Val Val Phe Phe Leu Leu Ser Ile Leu
1   5   10   15

Phe Ile Leu Tyr Asn Ile Ser Thr Val Cys His Tyr Tyr Met Arg Ile
            20   25   30

Ser Phe Tyr Tyr Phe Thr Ile Leu Leu His Gly Met Glu Val Cys Val
        35   40   45

Thr Met Ile Pro Ser Trp Leu Asn Gly Lys Gly Ala Asp Tyr Val Phe
    50   55   60

His Ser Phe Phe Tyr Trp Cys Lys Trp Thr Gly Val His Thr Thr Val
65   70   75   80

Tyr Gly Tyr Glu Lys Thr Gln Val Glu Gly Pro Ala Val Val Ile Cys
            85   90   95

Asn His Gln Ser Ser Leu Asp Ile Leu Ser Met Ala Ser Ile Trp Pro
        100   105   110

Lys Asn Cys Val Val Met Met Lys Arg Ile Leu Ala Tyr Val Pro Phe
        115   120   125

Phe Asn Leu Gly Ala Tyr Phe Ser Asn Thr Ile Phe Ile Asp Arg Tyr
    130   135   140

Asn Arg Glu Arg Ala Met Ala Ser Val Asp Tyr Cys Ala Ser Glu Met
145   150   155   160

Lys Asn Arg Asn Leu Lys Leu Trp Val Phe Pro Glu Gly Thr Arg Asn
                165   170   175

Arg Glu Gly Gly Phe Ile Pro Phe Lys Lys Gly Ala Phe Asn Ile Ala
            180   185   190

Val Arg Ala Gln Ile Pro Ile Ile Pro Val Val Phe Ser Asp Tyr Arg
        195   200   205

Asp Phe Tyr Ser Lys Pro Gly Arg Tyr Phe Lys Asn Asp Gly Glu Val
    210   215   220

```
Val Ile Arg Val Leu Asp Ala Ile Pro Thr Lys Gly Leu Thr Leu Asp
225                 230             235             240


Asp Val Ser Glu Leu Ser Asp Met Cys Arg Asp Val Met Leu Ala Ala
            245             250             255


Tyr Lys Glu Val Thr Leu Glu Ala Gln Gln Arg Asn Ala Thr Arg Arg
        260             265             270


Gly Glu Thr Lys Asp Gly Lys Lys Ser Glu
        275             280
```

```
<210>  3
<211>  849
<212>  DNA
<213>  Caenorhabditis elegans

<220>
<221>  CDS
<222>  (1)...(849)
<223>  Acyl-CoA:lysophospholipid acyltransferase


<400>  3
atg gag aac ttc tgg tcg atc gtc gtg ttt ttt cta ctc tca att ctc      48
Met Glu Asn Phe Trp Ser Ile Val Val Phe Phe Leu Leu Ser Ile Leu
1               5                   10                  15

ttc att tta tat aac ata tcg aca gta tgc cac tac tat atg cgg att      96
Phe Ile Leu Tyr Asn Ile Ser Thr Val Cys His Tyr Tyr Met Arg Ile
            20                  25                  30

tcg ttt tat tac ttc aca att tta ttg cat gga atg gaa gtt tgt gtt     144
Ser Phe Tyr Tyr Phe Thr Ile Leu Leu His Gly Met Glu Val Cys Val
        35                  40                  45

aca atg atc cct tct tgg cta aat ggg aag ggt gct gat tac gtg ttt     192
Thr Met Ile Pro Ser Trp Leu Asn Gly Lys Gly Ala Asp Tyr Val Phe
    50                  55                  60

cac tcg ttt ttc tat tgg tgt aaa tgg act ggt gtt cat aca aca gtc     240
His Ser Phe Phe Tyr Trp Cys Lys Trp Thr Gly Val His Thr Thr Val
65                  70                  75                  80

tat gga tat gaa aaa aca caa gtt gaa ggt ccg gct gta gtt att tgt     288
Tyr Gly Tyr Glu Lys Thr Gln Val Glu Gly Pro Ala Val Val Ile Cys
                85                  90                  95

aat cat cag agt tct ctc gac att cta tcg atg gca tca atc tgg ccg     336
Asn His Gln Ser Ser Leu Asp Ile Leu Ser Met Ala Ser Ile Trp Pro
            100                 105                 110

aag aat tgt gtt gta atg atg aaa cga att ctt gcc tat gtt cca ttc     384
Lys Asn Cys Val Val Met Met Lys Arg Ile Leu Ala Tyr Val Pro Phe
```

```
Lys Asn Cys Val Val Met Met Lys Arg Ile Leu Ala Tyr Val Pro Phe
    115                 120                 125

ttc aat ctc gga gcc tac ttt tcc aac aca atc ttc atc gat cga tat          432
Phe Asn Leu Gly Ala Tyr Phe Ser Asn Thr Ile Phe Ile Asp Arg Tyr
    130                 135                 140

aac cgt gaa cgt gcg atg gct tca gtt gat tat tgt gca tct gaa atg          480
Asn Arg Glu Arg Ala Met Ala Ser Val Asp Tyr Cys Ala Ser Glu Met
145                 150                 155                 160

aag aac aga aat ctt aaa ctt tgg gta tct ccg gaa gga aca aga aat          528
Lys Asn Arg Asn Leu Lys Leu Trp Val Ser Pro Glu Gly Thr Arg Asn
                165                 170                 175

cgt gaa gga ggg ttc att cca ttc aag aaa gga gca ttc aat att gca          576
Arg Glu Gly Gly Phe Ile Pro Phe Lys Lys Gly Ala Phe Asn Ile Ala
            180                 185                 190

gtt cgt gcg cag att ccc att att cca gtt gta ttc tca gac tat cgg          624
Val Arg Ala Gln Ile Pro Ile Ile Pro Val Val Phe Ser Asp Tyr Arg
            195                 200                 205

gat ttc tac tca aag cca ggc cga tat ttc aag aat gat gga gaa gtt          672
Asp Phe Tyr Ser Lys Pro Gly Arg Tyr Phe Lys Asn Asp Gly Glu Val
    210                 215                 220

gtt att cga gtt ctg gat gcg att cca aca aaa ggg ctc act ctt gat          720
Val Ile Arg Val Leu Asp Ala Ile Pro Thr Lys Gly Leu Thr Leu Asp
225                 230                 235                 240

gac gtc agc gag ttg tct gat atg tgt cgg gac gtt atg ttg gca gcc          768
Asp Val Ser Glu Leu Ser Asp Met Cys Arg Asp Val Met Leu Ala Ala
                245                 250                 255

tat aag gaa gtt act cta gaa gct cag caa cga aat gcg aca cgg cgt          816
Tyr Lys Glu Val Thr Leu Glu Ala Gln Gln Arg Asn Ala Thr Arg Arg
                260                 265                 270

gga gaa aca aaa gac ggg aag aaa tct gag taa                              849
Gly Glu Thr Lys Asp Gly Lys Lys Ser Glu
            275                 280
```

```
<210>   4
<211>   282
<212>   PRT
<213>   Caenorhabditis elegans

<400>   4

Met Glu Asn Phe Trp Ser Ile Val Val Phe Phe Leu Leu Ser Ile Leu
1                   5                   10                  15


Phe Ile Leu Tyr Asn Ile Ser Thr Val Cys His Tyr Tyr Met Arg Ile
            20                  25                  30
```

Ser Phe Tyr Tyr Phe Thr Ile Leu Leu His Gly Met Glu Val Cys Val
        35   40   45

Thr Met Ile Pro Ser Trp Leu Asn Gly Lys Gly Ala Asp Tyr Val Phe
        50   55   60

His Ser Phe Phe Tyr Trp Cys Lys Trp Thr Gly Val His Thr Thr Val
65   70   75   80

Tyr Gly Tyr Glu Lys Thr Gln Val Glu Gly Pro Ala Val Val Ile Cys
                85   90   95

Asn His Gln Ser Ser Leu Asp Ile Leu Ser Met Ala Ser Ile Trp Pro
            100   105   110

Lys Asn Cys Val Val Met Met Lys Arg Ile Leu Ala Tyr Val Pro Phe
        115   120   125

Phe Asn Leu Gly Ala Tyr Phe Ser Asn Thr Ile Phe Ile Asp Arg Tyr
        130   135   140

Asn Arg Glu Arg Ala Met Ala Ser Val Asp Tyr Cys Ala Ser Glu Met
145   150   155   160

Lys Asn Arg Asn Leu Lys Leu Trp Val Ser Pro Glu Gly Thr Arg Asn
            165   170   175

Arg Glu Gly Gly Phe Ile Pro Phe Lys Lys Gly Ala Phe Asn Ile Ala
            180   185   190

Val Arg Ala Gln Ile Pro Ile Ile Pro Val Val Phe Ser Asp Tyr Arg
        195   200   205

Asp Phe Tyr Ser Lys Pro Gly Arg Tyr Phe Lys Asn Asp Gly Glu Val
        210   215   220

Val Ile Arg Val Leu Asp Ala Ile Pro Thr Lys Gly Leu Thr Leu Asp
225   230   235   240

Asp Val Ser Glu Leu Ser Asp Met Cys Arg Asp Val Met Leu Ala Ala
            245   250   255

```
Tyr Lys Glu Val Thr Leu Glu Ala Gln Gln Arg Asn Ala Thr Arg Arg
        260                 265                 270

Gly Glu Thr Lys Asp Gly Lys Lys Ser Glu
        275                 280
```

<210>   5
<211>   849
<212>   DNA
<213>   Caenorhabditis elegans

<220>
<221>   CDS
<222>   (1)..(849)
<223>   Acyl-CoA:lysophospholipid acyltransferase

<400>   5

```
atg gag aac ttc tgg tcg atc gtc gtg ttt ttt cta ctc tca att ctc        48
Met Glu Asn Phe Trp Ser Ile Val Val Phe Phe Leu Leu Ser Ile Leu
1               5                   10                  15

ttc att tta tat aac ata tcg aca gta tgc cac tac tat gtg cgg att        96
Phe Ile Leu Tyr Asn Ile Ser Thr Val Cys His Tyr Tyr Val Arg Ile
            20                  25                  30

tcg ttt tat tac ttc aca att tta ttg cat gga atg gaa gtt tgt gtt       144
Ser Phe Tyr Tyr Phe Thr Ile Leu Leu His Gly Met Glu Val Cys Val
        35                  40                  45

aca atg atc cct tct tgg cta aat ggg aag ggt gct gat tac gtg ttt       192
Thr Met Ile Pro Ser Trp Leu Asn Gly Lys Gly Ala Asp Tyr Val Phe
    50                  55                  60

cac tcg ttt ttc tat tgg tgt aaa tgg act ggt gtt cat aca aca gtc       240
His Ser Phe Phe Tyr Trp Cys Lys Trp Thr Gly Val His Thr Thr Val
65                  70                  75                  80

tat gga tat gaa aaa aca caa gtt gaa ggt ccg gct gta gtt att tgt       288
Tyr Gly Tyr Glu Lys Thr Gln Val Glu Gly Pro Ala Val Val Ile Cys
                85                  90                  95

aat cat cag agt tct ctc gac att cta tcg atg gca tca atc tgg ccg       336
Asn His Gln Ser Ser Leu Asp Ile Leu Ser Met Ala Ser Ile Trp Pro
            100                 105                 110

aag aat tgt gtt gta atg atg aaa cga att ctt gcc tat gtt cca ttc       384
Lys Asn Cys Val Val Met Met Lys Arg Ile Leu Ala Tyr Val Pro Phe
        115                 120                 125

ttc aat ctc gga gcc tac ttt tcc aac aca atc ttc atc gat cga tat       432
Phe Asn Leu Gly Ala Tyr Phe Ser Asn Thr Ile Phe Ile Asp Arg Tyr
    130                 135                 140
```

```
aac cgt gaa cgt gcg atg gct tca gtt gat tat tgt gca tct gaa atg       480
Asn Arg Glu Arg Ala Met Ala Ser Val Asp Tyr Cys Ala Ser Glu Met
145             150             155             160

aag aac aga aat ctt aaa ctt tgg gta ttt ccg gaa gga aca aga aat       528
Lys Asn Arg Asn Leu Lys Leu Trp Val Phe Pro Glu Gly Thr Arg Asn
                165             170             175

cgt gaa gga ggg ttc att cca ttc aag aaa gga gca ttc aat att gca       576
Arg Glu Gly Gly Phe Ile Pro Phe Lys Lys Gly Ala Phe Asn Ile Ala
            180             185             190

gtt cgt gcg cag att ccc att att cca gtt gta ttc tca gac tat cgg       624
Val Arg Ala Gln Ile Pro Ile Ile Pro Val Val Phe Ser Asp Tyr Arg
            195             200             205

gat ttc tac tca aag cca ggc cga tat ttc aag aat gat gga gaa gtt       672
Asp Phe Tyr Ser Lys Pro Gly Arg Tyr Phe Lys Asn Asp Gly Glu Val
            210             215             220

gtt att cga gtt ctg gat gcg att cca aca aaa ggg ctc act ctt gat       720
Val Ile Arg Val Leu Asp Ala Ile Pro Thr Lys Gly Leu Thr Leu Asp
225             230             235             240

gac gtc agc gag ttg tct gat atg tgt cgg gac gtt atg ttg gca gcc       768
Asp Val Ser Glu Leu Ser Asp Met Cys Arg Asp Val Met Leu Ala Ala
                245             250             255

tat aag gaa gtt act cta gaa gct cag caa cga aat gcg aca cgg cgt       816
Tyr Lys Glu Val Thr Leu Glu Ala Gln Gln Arg Asn Ala Thr Arg Arg
                260             265             270

gga gaa aca aaa gac ggg aag aaa tct gag taa                            849
Gly Glu Thr Lys Asp Gly Lys Lys Ser Glu
            275             280
```

<210> 6
<211> 282
<212> PRT
<213> Caenorhabditis elegans

<400> 6

```
Met Glu Asn Phe Trp Ser Ile Val Val Phe Phe Leu Leu Ser Ile Leu
1               5               10              15


Phe Ile Leu Tyr Asn Ile Ser Thr Val Cys His Tyr Tyr Val Arg Ile
            20              25              30


Ser Phe Tyr Tyr Phe Thr Ile Leu Leu His Gly Met Glu Val Cys Val
        35              40              45


Thr Met Ile Pro Ser Trp Leu Asn Gly Lys Gly Ala Asp Tyr Val Phe
```

```
         50                        55                        60

His Ser Phe Phe Tyr Trp Cys Lys Trp Thr Gly Val His Thr Thr Val
65                  70                  75                  80

Tyr Gly Tyr Glu Lys Thr Gln Val Glu Gly Pro Ala Val Val Ile Cys
                85                  90                      95

Asn His Gln Ser Ser Leu Asp Ile Leu Ser Met Ala Ser Ile Trp Pro
                100                 105                 110

Lys Asn Cys Val Val Met Met Lys Arg Ile Leu Ala Tyr Val Pro Phe
            115                 120                 125

Phe Asn Leu Gly Ala Tyr Phe Ser Asn Thr Ile Phe Ile Asp Arg Tyr
    130                 135                 140

Asn Arg Glu Arg Ala Met Ala Ser Val Asp Tyr Cys Ala Ser Glu Met
145                 150                 155                 160

Lys Asn Arg Asn Leu Lys Leu Trp Val Phe Pro Glu Gly Thr Arg Asn
                165                 170                 175

Arg Glu Gly Gly Phe Ile Pro Phe Lys Lys Gly Ala Phe Asn Ile Ala
                180                 185                 190

Val Arg Ala Gln Ile Pro Ile Ile Pro Val Val Phe Ser Asp Tyr Arg
        195                 200                 205

Asp Phe Tyr Ser Lys Pro Gly Arg Tyr Phe Lys Asn Asp Gly Glu Val
    210                 215                 220

Val Ile Arg Val Leu Asp Ala Ile Pro Thr Lys Gly Leu Thr Leu Asp
225                 230                 235                 240

Asp Val Ser Glu Leu Ser Asp Met Cys Arg Asp Val Met Leu Ala Ala
                245                 250                 255

Tyr Lys Glu Val Thr Leu Glu Ala Gln Gln Arg Asn Ala Thr Arg Arg
                260                 265                 270

Gly Glu Thr Lys Asp Gly Lys Lys Ser Glu
                275                 280
```

```
<210>   7
<211>   849
<212>   DNA
<213>   Caenorhabditis elegans

<220>
<221>   CDS
<222>   (1)..(849)
<223>   Acyl-CoA:lysophospholipid acyltransferase


<400>   7
atg gag aac ttc tgg tcg atc gtc gtg ttt ttt cta ctc tca att ctc        48
Met Glu Asn Phe Trp Ser Ile Val Val Phe Phe Leu Leu Ser Ile Leu
1               5                   10                  15

ttc att tta tat aac ata tcg aca gta tgc cac tac tat atg cgg att        96
Phe Ile Leu Tyr Asn Ile Ser Thr Val Cys His Tyr Tyr Met Arg Ile
                20                  25                  30

tcg ttt tat tac ttc aca att tta ttg cat gga atg gaa gtt tgt gtt       144
Ser Phe Tyr Tyr Phe Thr Ile Leu Leu His Gly Met Glu Val Cys Val
        35                  40                  45

aca atg atc cct tct tgg cta aat ggg aag ggt gct gat tac gtg ttt       192
Thr Met Ile Pro Ser Trp Leu Asn Gly Lys Gly Ala Asp Tyr Val Phe
    50                  55                  60

cac tcg ttt ttc tat tgg tgt aaa tgg act ggt gtt cat aca aca gtc       240
His Ser Phe Phe Tyr Trp Cys Lys Trp Thr Gly Val His Thr Thr Val
65                  70                  75                  80

tat gga tat gaa aaa aca caa gtt gaa ggt ccg gcc gta gtt att tgt       288
Tyr Gly Tyr Glu Lys Thr Gln Val Glu Gly Pro Ala Val Val Ile Cys
                85                  90                  95

aat cat cag ggt tct ctc gac att cta tcg atg gca tca atc tgg ccg       336
Asn His Gln Gly Ser Leu Asp Ile Leu Ser Met Ala Ser Ile Trp Pro
                100                 105                 110

aag aat tgt gtt gta atg atg aaa cga att ctt gcc tat gtt cca ttc       384
Lys Asn Cys Val Val Met Met Lys Arg Ile Leu Ala Tyr Val Pro Phe
            115                 120                 125

ttc aat ctc gga gcc tac ttt tcc aac aca atc ttc atc gat cga tat       432
Phe Asn Leu Gly Ala Tyr Phe Ser Asn Thr Ile Phe Ile Asp Arg Tyr
    130                 135                 140

aac cgt gaa cgt gcg atg gct tca gtt gat tat tgt gca tct gaa atg       480
Asn Arg Glu Arg Ala Met Ala Ser Val Asp Tyr Cys Ala Ser Glu Met
145                 150                 155                 160

aag aac aga aat ctt aaa ctt tgg gta ttt ccg gaa gga aca aga aat       528
Lys Asn Arg Asn Leu Lys Leu Trp Val Phe Pro Glu Gly Thr Arg Asn
                165                 170                 175
```

```
cgt gaa gga ggg ttc att cca ttc aag aaa gga gca ttc aat att gca        576
Arg Glu Gly Gly Phe Ile Pro Phe Lys Lys Gly Ala Phe Asn Ile Ala
            180             185             190

gtt cgt gcg cag att ccc att att cca gtt gta ttc tca gac tat cgg        624
Val Arg Ala Gln Ile Pro Ile Ile Pro Val Val Phe Ser Asp Tyr Arg
        195             200             205

gat ttc tac tca aag cca ggc cga tat ttc aag aat gat gga gaa gtt        672
Asp Phe Tyr Ser Lys Pro Gly Arg Tyr Phe Lys Asn Asp Gly Glu Val
    210             215             220

gtt att cga gtt ctg gat gcg att cca aca aaa ggg ctc act ctt gat        720
Val Ile Arg Val Leu Asp Ala Ile Pro Thr Lys Gly Leu Thr Leu Asp
225             230             235             240

gac gtc agc gag ttg tct gat atg tgt cgg gac gtt atg ttg gca gcc        768
Asp Val Ser Glu Leu Ser Asp Met Cys Arg Asp Val Met Leu Ala Ala
            245             250             255

tat aag gaa gtt act cta gaa gct cag caa cga aat gcg aca cgg cgt        816
Tyr Lys Glu Val Thr Leu Glu Ala Gln Gln Arg Asn Ala Thr Arg Arg
            260             265             270

gga gaa aca aaa gac ggg aag aaa tct gag taa                            849
Gly Glu Thr Lys Asp Gly Lys Lys Ser Glu
        275             280
```

<210>   8
<211>   282
<212>   PRT
<213>   Caenorhabditis elegans

<400>   8

```
Met Glu Asn Phe Trp Ser Ile Val Val Phe Phe Leu Leu Ser Ile Leu
1               5               10              15


Phe Ile Leu Tyr Asn Ile Ser Thr Val Cys His Tyr Tyr Met Arg Ile
            20              25              30


Ser Phe Tyr Tyr Phe Thr Ile Leu Leu His Gly Met Glu Val Cys Val
        35              40              45


Thr Met Ile Pro Ser Trp Leu Asn Gly Lys Gly Ala Asp Tyr Val Phe
    50              55              60


His Ser Phe Phe Tyr Trp Cys Lys Trp Thr Gly Val His Thr Thr Val
65              70              75              80
```

Tyr Gly Tyr Glu Lys Thr Gln Val Glu Gly Pro Ala Val Val Ile Cys
            85   90   95

Asn His Gln Gly Ser Leu Asp Ile Leu Ser Met Ala Ser Ile Trp Pro
           100   105   110

Lys Asn Cys Val Val Met Met Lys Arg Ile Leu Ala Tyr Val Pro Phe
           115   120   125

Phe Asn Leu Gly Ala Tyr Phe Ser Asn Thr Ile Phe Ile Asp Arg Tyr
   130   135   140

Asn Arg Glu Arg Ala Met Ala Ser Val Asp Tyr Cys Ala Ser Glu Met
145   150   155   160

Lys Asn Arg Asn Leu Lys Leu Trp Val Phe Pro Glu Gly Thr Arg Asn
           165   170   175

Arg Glu Gly Gly Phe Ile Pro Phe Lys Lys Gly Ala Phe Asn Ile Ala
           180   185   190

Val Arg Ala Gln Ile Pro Ile Ile Pro Val Val Phe Ser Asp Tyr Arg
          195   200   205

Asp Phe Tyr Ser Lys Pro Gly Arg Tyr Phe Lys Asn Asp Gly Glu Val
   210   215   220

Val Ile Arg Val Leu Asp Ala Ile Pro Thr Lys Gly Leu Thr Leu Asp
225   230   235   240

Asp Val Ser Glu Leu Ser Asp Met Cys Arg Asp Val Met Leu Ala Ala
           245   250   255

Tyr Lys Glu Val Thr Leu Glu Ala Gln Gln Arg Asn Ala Thr Arg Arg
           260   265   270

Gly Glu Thr Lys Asp Gly Lys Lys Ser Glu
          275   280

<210>   9
<211>   1578
<212>   DNA
<213>   Physcomitrella patens

```
<220>
<221>  CDS
<222>  (1)..(1578)
<223>  Delta-6-desaturase


<400>  9
atg gta ttc gcg ggc ggt gga ctt cag cag ggc tct ctc gaa gaa aac      48
Met Val Phe Ala Gly Gly Gly Leu Gln Gln Gly Ser Leu Glu Glu Asn
1               5                   10                  15

atc gac gtc gag cac att gcc agt atg tct ctc ttc agc gac ttc ttc      96
Ile Asp Val Glu His Ile Ala Ser Met Ser Leu Phe Ser Asp Phe Phe
                20                  25                  30

agt tat gtg tct tca act gtt ggt tcg tgg agc gta cac agt ata caa     144
Ser Tyr Val Ser Ser Thr Val Gly Ser Trp Ser Val His Ser Ile Gln
            35                  40                  45

cct ttg aag cgc ctg acg agt aag aag cgt gtt tcg gaa agc gct gcc     192
Pro Leu Lys Arg Leu Thr Ser Lys Lys Arg Val Ser Glu Ser Ala Ala
        50                  55                  60

gtg caa tgt ata tca gct gaa gtt cag aga aat tcg agt acc cag gga     240
Val Gln Cys Ile Ser Ala Glu Val Gln Arg Asn Ser Ser Thr Gln Gly
65                  70                  75                  80

act gcg gag gca ctc gca gaa tca gtc gtg aag ccc acg aga cga agg     288
Thr Ala Glu Ala Leu Ala Glu Ser Val Val Lys Pro Thr Arg Arg Arg
                85                  90                  95

tca tct cag tgg aag aag tcg aca cac ccc cta tca gaa gta gca gta     336
Ser Ser Gln Trp Lys Lys Ser Thr His Pro Leu Ser Glu Val Ala Val
                100                 105                 110

cac aac aag cca agc gat tgc tgg att gtt gta aaa aac aag gtg tat     384
His Asn Lys Pro Ser Asp Cys Trp Ile Val Val Lys Asn Lys Val Tyr
            115                 120                 125

gat gtt tcc aat ttt gcg gac gag cat ccc gga gga tca gtt att agt     432
Asp Val Ser Asn Phe Ala Asp Glu His Pro Gly Gly Ser Val Ile Ser
            130                 135                 140

act tat ttt gga cga gac ggc aca gat gtt ttc tct agt ttt cat gca     480
Thr Tyr Phe Gly Arg Asp Gly Thr Asp Val Phe Ser Ser Phe His Ala
145                 150                 155                 160

gct tct aca tgg aaa att ctt caa gac ttt tac att ggt gac gtg gag     528
Ala Ser Thr Trp Lys Ile Leu Gln Asp Phe Tyr Ile Gly Asp Val Glu
                165                 170                 175

agg gtg gag ccg act cca gag ctg ctg aaa gat ttc cga gaa atg aga     576
Arg Val Glu Pro Thr Pro Glu Leu Leu Lys Asp Phe Arg Glu Met Arg
                180                 185                 190

gct ctt ttc ctg agg gag caa ctt ttc aaa agt tcg aaa ttg tac tat     624
```

```
Ala Leu Phe Leu Arg Glu Gln Leu Phe Lys Ser Ser Lys Leu Tyr Tyr
    195                 200                 205

gtt atg aag ctg ctc acg aat gtt gct att ttt gct gcg agc att gca    672
Val Met Lys Leu Leu Thr Asn Val Ala Ile Phe Ala Ala Ser Ile Ala
    210             215                 220

ata ata tgt tgg agc aag act att tca gcg gtt ttg gct tca gct tgt    720
Ile Ile Cys Trp Ser Lys Thr Ile Ser Ala Val Leu Ala Ser Ala Cys
225             230             235                 240

atg atg gct ctg tgt ttc caa cag tgc gga tgg cta tcc cat gat ttt    768
Met Met Ala Leu Cys Phe Gln Gln Cys Gly Trp Leu Ser His Asp Phe
            245             250             255

ctc cac aat cag gtg ttt gag aca cgc tgg ctt aat gaa gtt gtc ggg    816
Leu His Asn Gln Val Phe Glu Thr Arg Trp Leu Asn Glu Val Val Gly
            260             265             270

tat gtg atc ggc aac gcc gtt ctg ggg ttt agt aca ggg tgg tgg aag    864
Tyr Val Ile Gly Asn Ala Val Leu Gly Phe Ser Thr Gly Trp Trp Lys
            275             280             285

gag aag cat aac ctt cat cat gct gct cca aat gaa tgc gat cag act    912
Glu Lys His Asn Leu His His Ala Ala Pro Asn Glu Cys Asp Gln Thr
    290             295             300

tac caa cca att gat gaa gat att gat act ctc ccc ctc att gcc tgg    960
Tyr Gln Pro Ile Asp Glu Asp Ile Asp Thr Leu Pro Leu Ile Ala Trp
305             310             315             320

agc aag gac ata ctg gcc aca gtt gag aat aag aca ttc ttg cga atc    1008
Ser Lys Asp Ile Leu Ala Thr Val Glu Asn Lys Thr Phe Leu Arg Ile
            325             330             335

ctc caa tac cag cat ctg ttc ttc atg ggt ctg tta ttt ttc gcc cgt    1056
Leu Gln Tyr Gln His Leu Phe Phe Met Gly Leu Leu Phe Phe Ala Arg
            340             345             350

ggt agt tgg ctc ttt tgg agc tgg aga tat acc tct aca gca gtg ctc    1104
Gly Ser Trp Leu Phe Trp Ser Trp Arg Tyr Thr Ser Thr Ala Val Leu
            355             360             365

tca cct gtc gac agg ttg ttg gag aag gga act gtt ctg ttt cac tac    1152
Ser Pro Val Asp Arg Leu Leu Glu Lys Gly Thr Val Leu Phe His Tyr
            370             375             380

ttt tgg ttc gtc ggg aca gcg tgc tat ctt ctc cct ggt tgg aag cca    1200
Phe Trp Phe Val Gly Thr Ala Cys Tyr Leu Leu Pro Gly Trp Lys Pro
385             390             395             400

tta gta tgg atg gcg gtg act gag ctc atg tcc ggc atg ctg ctg ggc    1248
Leu Val Trp Met Ala Val Thr Glu Leu Met Ser Gly Met Leu Leu Gly
            405             410             415

ttt gta ttt gta ctt agc cac aat ggg atg gag gtt tat aat tcg tct    1296
Phe Val Phe Val Leu Ser His Asn Gly Met Glu Val Tyr Asn Ser Ser
```

```
                 420                      425                      430
aaa gaa ttc gtg agt gca cag atc gta tcc aca cgg gat atc aaa gga          1344
Lys Glu Phe Val Ser Ala Gln Ile Val Ser Thr Arg Asp Ile Lys Gly
        435                      440                      445

aac ata ttc aac gac tgg ttc act ggt ggc ctt aac agg caa ata gag          1392
Asn Ile Phe Asn Asp Trp Phe Thr Gly Gly Leu Asn Arg Gln Ile Glu
        450                      455                      460

cat cat ctt ttc cca aca atg ccc agg cat aat tta aac aaa ata gca          1440
His His Leu Phe Pro Thr Met Pro Arg His Asn Leu Asn Lys Ile Ala
465                      470                      475                      480

cct aga gtg gag gtg ttc tgt aag aaa cac ggt ctg gtg tac gaa gac          1488
Pro Arg Val Glu Val Phe Cys Lys Lys His Gly Leu Val Tyr Glu Asp
                 485                      490                      495

gta tct att gct acc ggc act tgc aag gtt ttg aaa gca ttg aag gaa          1536
Val Ser Ile Ala Thr Gly Thr Cys Lys Val Leu Lys Ala Leu Lys Glu
                 500                      505                      510

gtc gcg gag gct gcg gca gag cag cat gct acc acc agt taa                  1578
Val Ala Glu Ala Ala Ala Glu Gln His Ala Thr Thr Ser
        515                      520                      525
```

```
<210>  10
<211>  525
<212>  PRT
<213>  Physcomitrella patens

<400>  10

Met Val Phe Ala Gly Gly Gly Leu Gln Gln Gly Ser Leu Glu Glu Asn
1               5                   10                  15

Ile Asp Val Glu His Ile Ala Ser Met Ser Leu Phe Ser Asp Phe Phe
             20                  25                  30

Ser Tyr Val Ser Ser Thr Val Gly Ser Trp Ser Val His Ser Ile Gln
        35                  40                  45

Pro Leu Lys Arg Leu Thr Ser Lys Lys Arg Val Ser Glu Ser Ala Ala
     50                  55                  60

Val Gln Cys Ile Ser Ala Glu Val Gln Arg Asn Ser Ser Thr Gln Gly
65                  70                  75                  80

Thr Ala Glu Ala Leu Ala Glu Ser Val Val Lys Pro Thr Arg Arg Arg
                 85                  90                  95
```

```
Ser Ser Gln Trp Lys Lys Ser Thr His Pro Leu Ser Glu Val Ala Val
        100             105               110

His Asn Lys Pro Ser Asp Cys Trp Ile Val Val Lys Asn Lys Val Tyr
        115             120               125

Asp Val Ser Asn Phe Ala Asp Glu His Pro Gly Gly Ser Val Ile Ser
    130             135             140

Thr Tyr Phe Gly Arg Asp Gly Thr Asp Val Phe Ser Ser Phe His Ala
145             150             155             160

Ala Ser Thr Trp Lys Ile Leu Gln Asp Phe Tyr Ile Gly Asp Val Glu
                165             170             175

Arg Val Glu Pro Thr Pro Glu Leu Leu Lys Asp Phe Arg Glu Met Arg
            180             185             190

Ala Leu Phe Leu Arg Glu Gln Leu Phe Lys Ser Ser Lys Leu Tyr Tyr
        195             200             205

Val Met Lys Leu Leu Thr Asn Val Ala Ile Phe Ala Ala Ser Ile Ala
    210             215             220

Ile Ile Cys Trp Ser Lys Thr Ile Ser Ala Val Leu Ala Ser Ala Cys
225             230             235             240

Met Met Ala Leu Cys Phe Gln Gln Cys Gly Trp Leu Ser His Asp Phe
                245             250             255

Leu His Asn Gln Val Phe Glu Thr Arg Trp Leu Asn Glu Val Val Gly
        260             265             270

Tyr Val Ile Gly Asn Ala Val Leu Gly Phe Ser Thr Gly Trp Trp Lys
        275             280             285

Glu Lys His Asn Leu His His Ala Ala Pro Asn Glu Cys Asp Gln Thr
    290             295             300

Tyr Gln Pro Ile Asp Glu Asp Ile Asp Thr Leu Pro Leu Ile Ala Trp
305             310             315             320
```

Ser Lys Asp Ile Leu Ala Thr Val Glu Asn Lys Thr Phe Leu Arg Ile
            325   330   335

Leu Gln Tyr Gln His Leu Phe Phe Met Gly Leu Leu Phe Phe Ala Arg
            340   345   350

Gly Ser Trp Leu Phe Trp Ser Trp Arg Tyr Thr Ser Thr Ala Val Leu
            355   360   365

Ser Pro Val Asp Arg Leu Leu Glu Lys Gly Thr Val Leu Phe His Tyr
    370   375   380

Phe Trp Phe Val Gly Thr Ala Cys Tyr Leu Leu Pro Gly Trp Lys Pro
385   390   395   400

Leu Val Trp Met Ala Val Thr Glu Leu Met Ser Gly Met Leu Leu Gly
            405   410   415

Phe Val Phe Val Leu Ser His Asn Gly Met Glu Val Tyr Asn Ser Ser
            420   425   430

Lys Glu Phe Val Ser Ala Gln Ile Val Ser Thr Arg Asp Ile Lys Gly
            435   440   445

Asn Ile Phe Asn Asp Trp Phe Thr Gly Gly Leu Asn Arg Gln Ile Glu
    450   455   460

His His Leu Phe Pro Thr Met Pro Arg His Asn Leu Asn Lys Ile Ala
465   470   475   480

Pro Arg Val Glu Val Phe Cys Lys Lys His Gly Leu Val Tyr Glu Asp
            485   490   495

Val Ser Ile Ala Thr Gly Thr Cys Lys Val Leu Lys Ala Leu Lys Glu
            500   505   510

Val Ala Glu Ala Ala Ala Glu Gln His Ala Thr Thr Ser
            515   520   525

<210>   11
<211>   1192
<212>   DNA
<213>   Physcomitrella patens

```
<220>
<221>  CDS
<222>  (58)..(930)
<223>  Delta-6-elongase


<400>  11
ctgcttcgtc tcatcttggg ggtgtgattc gggagtgggt tgagttggtg gagcgca          57

atg gag gtc gtg gag aga ttc tac ggt gag ttg gat ggg aag gtc tcg     105
Met Glu Val Val Glu Arg Phe Tyr Gly Glu Leu Asp Gly Lys Val Ser
1               5                   10                  15

cag ggc gtg aat gca ttg ctg ggt agt ttt ggg gtg gag ttg acg gat     153
Gln Gly Val Asn Ala Leu Leu Gly Ser Phe Gly Val Glu Leu Thr Asp
            20                  25                  30

acg ccc act acc aaa ggc ttg ccc ctc gtt gac agt ccc aca ccc atc     201
Thr Pro Thr Thr Lys Gly Leu Pro Leu Val Asp Ser Pro Thr Pro Ile
        35                  40                  45

gtc ctc ggt gtt tct gta tac ttg act att gtc att gga ggg ctt ttg     249
Val Leu Gly Val Ser Val Tyr Leu Thr Ile Val Ile Gly Gly Leu Leu
    50                  55                  60

tgg ata aag gcc agg gat ctg aaa ccg cgc gcc tcg gag cca ttt ttg     297
Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg Ala Ser Glu Pro Phe Leu
65                  70                  75                  80

ctc caa gct ttg gtg ctt gtg cac aac ctg ttc tgt ttt gcg ctc agt     345
Leu Gln Ala Leu Val Leu Val His Asn Leu Phe Cys Phe Ala Leu Ser
                85                  90                  95

ctg tat atg tgc gtg ggc atc gct tat cag gct att acc tgg cgg tac     393
Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln Ala Ile Thr Trp Arg Tyr
                100                 105                 110

tct ctc tgg ggc aat gca tac aat cct aaa cat aaa gag atg gcg att     441
Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys His Lys Glu Met Ala Ile
        115                 120                 125

ctg gta tac ttg ttc tac atg tct aag tac gtg gaa ttc atg gat acc     489
Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr Val Glu Phe Met Asp Thr
    130                 135                 140

gtt atc atg ata ctg aag cgc agc acc agg caa ata agc ttc ctc cac     537
Val Ile Met Ile Leu Lys Arg Ser Thr Arg Gln Ile Ser Phe Leu His
145                 150                 155                 160

gtt tat cat cat tct tca att tcc ctc att tgg tgg gct att gct cat     585
Val Tyr His His Ser Ser Ile Ser Leu Ile Trp Trp Ala Ile Ala His
                165                 170                 175

cac gct cct ggc ggt gaa gca tat tgg tct gcg gct ctg aac tca gga     633
His Ala Pro Gly Gly Glu Ala Tyr Trp Ser Ala Ala Leu Asn Ser Gly
                180                 185                 190
```

```
gtg cat gtt ctc atg tat gcg tat tac ttc ttg gct gcc tgc ctt cga      681
Val His Val Leu Met Tyr Ala Tyr Tyr Phe Leu Ala Ala Cys Leu Arg
        195                 200                 205

agt agc cca aag tta aaa aat aag tac ctt ttt tgg ggc agg tac ttg      729
Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu Phe Trp Gly Arg Tyr Leu
        210                 215                 220

aca caa ttc caa atg ttc cag ttt atg ctg aac tta gtg cag gct tac      777
Thr Gln Phe Gln Met Phe Gln Phe Met Leu Asn Leu Val Gln Ala Tyr
225                 230                 235                 240

tac gac atg aaa acg aat gcg cca tat cca caa tgg ctg atc aag att      825
Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro Gln Trp Leu Ile Lys Ile
                245                 250                 255

ttg ttc tac tac atg atc tcg ttg ctg ttt ctt ttc ggc aat ttt tac      873
Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe Leu Phe Gly Asn Phe Tyr
                260                 265                 270

gta caa aaa tac atc aaa ccc tct gac gga aag caa aag gga gct aaa      921
Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly Lys Gln Lys Gly Ala Lys
        275                 280                 285

act gag tga gctgtatcaa gccatagaaa ctctattatg ttagaacctg            970
Thr Glu
    290

aagttggtgc tttcttatct ccacttatct tttaagcagc atcagttttg aaatgatgtg   1030

tgggcgtggt ctgcaagtag tcatcaatat aatcggcctg agcacttcag atggattgtt   1090

agaacatgag taaaagcggt tattacggtg tttattttgt accaaatcac cgcacgggtg   1150

aattgaaata tttcagattt gatcaatttc atctgaaaaa aa                      1192
```

```
<210>  12
<211>  290
<212>  PRT
<213>  Physcomitrella patens

<400>  12

Met Glu Val Val Glu Arg Phe Tyr Gly Glu Leu Asp Gly Lys Val Ser
1               5               10                  15


Gln Gly Val Asn Ala Leu Leu Gly Ser Phe Gly Val Glu Leu Thr Asp
            20                  25                  30


Thr Pro Thr Thr Lys Gly Leu Pro Leu Val Asp Ser Pro Thr Pro Ile
        35                  40                  45


Val Leu Gly Val Ser Val Tyr Leu Thr Ile Val Ile Gly Gly Leu Leu
```

```
       50                      55                      60

Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg Ala Ser Glu Pro Phe Leu
65                  70              75                  80


Leu Gln Ala Leu Val Leu Val His Asn Leu Phe Cys Phe Ala Leu Ser
                85              90                  95


Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln Ala Ile Thr Trp Arg Tyr
            100             105                 110


Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys His Lys Glu Met Ala Ile
            115             120             125


Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr Val Glu Phe Met Asp Thr
            130             135             140


Val Ile Met Ile Leu Lys Arg Ser Thr Arg Gln Ile Ser Phe Leu His
145             150             155             160


Val Tyr His His Ser Ser Ile Ser Leu Ile Trp Trp Ala Ile Ala His
            165             170             175


His Ala Pro Gly Gly Glu Ala Tyr Trp Ser Ala Ala Leu Asn Ser Gly
            180             185             190


Val His Val Leu Met Tyr Ala Tyr Tyr Phe Leu Ala Ala Cys Leu Arg
            195             200             205


Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu Phe Trp Gly Arg Tyr Leu
    210             215             220


Thr Gln Phe Gln Met Phe Gln Phe Met Leu Asn Leu Val Gln Ala Tyr
225             230             235             240


Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro Gln Trp Leu Ile Lys Ile
            245             250             255


Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe Leu Phe Gly Asn Phe Tyr
            260             265             270


Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly Lys Gln Lys Gly Ala Lys
            275             280             285
```

Thr Glu
   290

<210> 13
<211> 1410
<212> DNA
<213> Phaeodactylum tricornutum

<220>
<221> CDS
<222> (1)..(1410)
<223> Delta-5-desaturase

<400> 13
```
atg gct ccg gat gcg gat aag ctt cga caa cgc cag acg act gcg gta          48
Met Ala Pro Asp Ala Asp Lys Leu Arg Gln Arg Gln Thr Thr Ala Val
1               5                   10                  15

gcg aag cac aat gct gct acc ata tcg acg cag gaa cgc ctt tgc agt          96
Ala Lys His Asn Ala Ala Thr Ile Ser Thr Gln Glu Arg Leu Cys Ser
                20                  25                  30

ctg tct tcg ctc aaa ggc gaa gaa gtc tgc atc gac gga atc atc tat         144
Leu Ser Ser Leu Lys Gly Glu Glu Val Cys Ile Asp Gly Ile Ile Tyr
            35                  40                  45

gac ctc caa tca ttc gat cat ccc ggg ggt gaa acg atc aaa atg ttt         192
Asp Leu Gln Ser Phe Asp His Pro Gly Gly Glu Thr Ile Lys Met Phe
        50                  55                  60

ggt ggc aac gat gtc act gta cag tac aag atg att cac ccg tac cat         240
Gly Gly Asn Asp Val Thr Val Gln Tyr Lys Met Ile His Pro Tyr His
65                  70                  75                  80

acc gag aag cat ttg gaa aag atg aag cgt gtc ggc aag gtg acg gat         288
Thr Glu Lys His Leu Glu Lys Met Lys Arg Val Gly Lys Val Thr Asp
                85                  90                  95

ttc gtc tgc gag tac aag ttc gat acc gaa ttt gaa cgc gaa atc aaa         336
Phe Val Cys Glu Tyr Lys Phe Asp Thr Glu Phe Glu Arg Glu Ile Lys
                100                 105                 110

cga gaa gtc ttc aag att gtg cga cga ggc aag gat ttc ggt act ttg         384
Arg Glu Val Phe Lys Ile Val Arg Arg Gly Lys Asp Phe Gly Thr Leu
            115                 120                 125

gga tgg ttc ttc cgt gcg ttt tgc tac att gcc att ttc ttc tac ctg         432
Gly Trp Phe Phe Arg Ala Phe Cys Tyr Ile Ala Ile Phe Phe Tyr Leu
        130                 135                 140

cag tac cat tgg gtc acc acg gga acc tct tgg ctg ctg gcc gtg gcc         480
Gln Tyr His Trp Val Thr Thr Gly Thr Ser Trp Leu Leu Ala Val Ala
```

```
    145                    150                    155                    160
tac gga atc tcc caa gcg atg att ggc atg aat gtc cag cac gat gcc        528
Tyr Gly Ile Ser Gln Ala Met Ile Gly Met Asn Val Gln His Asp Ala
            165                    170                    175

aac cac ggg gcc acc tcc aag cgt ccc tgg gtc aac gac atg cta ggc        576
Asn His Gly Ala Thr Ser Lys Arg Pro Trp Val Asn Asp Met Leu Gly
            180                    185                    190

ctc ggt gcg gat ttt att ggt ggt tcc aag tgg ctc tgg cag gaa caa        624
Leu Gly Ala Asp Phe Ile Gly Gly Ser Lys Trp Leu Trp Gln Glu Gln
            195                    200                    205

cac tgg acc cac cac gct tac acc aat cac gcc gag atg gat ccc gat        672
His Trp Thr His His Ala Tyr Thr Asn His Ala Glu Met Asp Pro Asp
    210                    215                    220

agc ttt ggt gcc gaa cca atg ctc cta ttc aac gac tat ccc ttg gat        720
Ser Phe Gly Ala Glu Pro Met Leu Leu Phe Asn Asp Tyr Pro Leu Asp
225                    230                    235                    240

cat ccc gct cgt acc tgg cta cat cgc ttt caa gca ttc ttt tac atg        768
His Pro Ala Arg Thr Trp Leu His Arg Phe Gln Ala Phe Phe Tyr Met
                245                    250                    255

ccc gtc ttg gct gga tac tgg ttg tcc gct gtc ttc aat cca caa att        816
Pro Val Leu Ala Gly Tyr Trp Leu Ser Ala Val Phe Asn Pro Gln Ile
            260                    265                    270

ctt gac ctc cag caa cgc ggc gca ctt tcc gtc ggt atc cgt ctc gac        864
Leu Asp Leu Gln Gln Arg Gly Ala Leu Ser Val Gly Ile Arg Leu Asp
            275                    280                    285

aac gct ttc att cac tcg cga cgc aag tat gcg gtt ttc tgg cgg gct        912
Asn Ala Phe Ile His Ser Arg Arg Lys Tyr Ala Val Phe Trp Arg Ala
            290                    295                    300

gtg tac att gcg gtg aac gtg att gct ccg ttt tac aca aac tcc ggc        960
Val Tyr Ile Ala Val Asn Val Ile Ala Pro Phe Tyr Thr Asn Ser Gly
305                    310                    315                    320

ctc gaa tgg tcc tgg cgt gtc ttt gga aac atc atg ctc atg ggt gtg       1008
Leu Glu Trp Ser Trp Arg Val Phe Gly Asn Ile Met Leu Met Gly Val
                325                    330                    335

gcg gaa tcg ctc gcg ctg gcg gtc ctg ttt tcg ttg tcg cac aat ttc       1056
Ala Glu Ser Leu Ala Leu Ala Val Leu Phe Ser Leu Ser His Asn Phe
            340                    345                    350

gaa tcc gcg gat cgc gat ccg acc gcc cca ctg aaa aag acg gga gaa       1104
Glu Ser Ala Asp Arg Asp Pro Thr Ala Pro Leu Lys Lys Thr Gly Glu
            355                    360                    365

cca gtc gac tgg ttc aag aca cag gtc gaa act tcc tgc act tac ggt       1152
Pro Val Asp Trp Phe Lys Thr Gln Val Glu Thr Ser Cys Thr Tyr Gly
            370                    375                    380
```

```
gga ttc ctt tcc ggt tgc ttc acg gga ggt ctc aac ttt cag gtt gaa      1200
Gly Phe Leu Ser Gly Cys Phe Thr Gly Gly Leu Asn Phe Gln Val Glu
385             390             395                 400

cac cac ttg ttc cca cgc atg agc agc gct tgg tat ccc tac att gcc      1248
His His Leu Phe Pro Arg Met Ser Ser Ala Trp Tyr Pro Tyr Ile Ala
            405             410             415

ccc aag gtc cgc gaa att tgc gcc aaa cac ggc gtc cac tac gcc tac      1296
Pro Lys Val Arg Glu Ile Cys Ala Lys His Gly Val His Tyr Ala Tyr
            420             425             430

tac ccg tgg atc cac caa aac ttt ctc tcc acc gtc cgc tac atg cac      1344
Tyr Pro Trp Ile His Gln Asn Phe Leu Ser Thr Val Arg Tyr Met His
            435             440             445

gcg gcc ggg acc ggt gcc aac tgg cgc cag atg gcc aga gaa aat ccc      1392
Ala Ala Gly Thr Gly Ala Asn Trp Arg Gln Met Ala Arg Glu Asn Pro
    450             455             460

ttg acc gga cgg gcg taa                                              1410
Leu Thr Gly Arg Ala
465


<210>   14
<211>   469
<212>   PRT
<213>   Phaeodactylum tricornutum

<400>   14

Met Ala Pro Asp Ala Asp Lys Leu Arg Gln Arg Gln Thr Thr Ala Val
1               5               10              15


Ala Lys His Asn Ala Ala Thr Ile Ser Thr Gln Glu Arg Leu Cys Ser
            20              25              30


Leu Ser Ser Leu Lys Gly Glu Glu Val Cys Ile Asp Gly Ile Ile Tyr
        35              40              45


Asp Leu Gln Ser Phe Asp His Pro Gly Gly Glu Thr Ile Lys Met Phe
    50              55              60


Gly Gly Asn Asp Val Thr Val Gln Tyr Lys Met Ile His Pro Tyr His
65              70              75                  80


Thr Glu Lys His Leu Glu Lys Met Lys Arg Val Gly Lys Val Thr Asp
                85              90              95
```

```
Phe Val Cys Glu Tyr Lys Phe Asp Thr Glu Phe Glu Arg Glu Ile Lys
            100             105             110

Arg Glu Val Phe Lys Ile Val Arg Arg Gly Lys Asp Phe Gly Thr Leu
            115             120             125

Gly Trp Phe Phe Arg Ala Phe Cys Tyr Ile Ala Ile Phe Phe Tyr Leu
    130             135             140

Gln Tyr His Trp Val Thr Thr Gly Thr Ser Trp Leu Leu Ala Val Ala
145             150             155             160

Tyr Gly Ile Ser Gln Ala Met Ile Gly Met Asn Val Gln His Asp Ala
            165             170             175

Asn His Gly Ala Thr Ser Lys Arg Pro Trp Val Asn Asp Met Leu Gly
            180             185             190

Leu Gly Ala Asp Phe Ile Gly Gly Ser Lys Trp Leu Trp Gln Glu Gln
            195             200             205

His Trp Thr His His Ala Tyr Thr Asn His Ala Glu Met Asp Pro Asp
    210             215             220

Ser Phe Gly Ala Glu Pro Met Leu Leu Phe Asn Asp Tyr Pro Leu Asp
225             230             235             240

His Pro Ala Arg Thr Trp Leu His Arg Phe Gln Ala Phe Phe Tyr Met
            245             250             255

Pro Val Leu Ala Gly Tyr Trp Leu Ser Ala Val Phe Asn Pro Gln Ile
            260             265             270

Leu Asp Leu Gln Gln Arg Gly Ala Leu Ser Val Gly Ile Arg Leu Asp
            275             280             285

Asn Ala Phe Ile His Ser Arg Arg Lys Tyr Ala Val Phe Trp Arg Ala
            290             295             300

Val Tyr Ile Ala Val Asn Val Ile Ala Pro Phe Tyr Thr Asn Ser Gly
305             310             315             320

Leu Glu Trp Ser Trp Arg Val Phe Gly Asn Ile Met Leu Met Gly Val
```

                    325   330   335

Ala Glu Ser Leu Ala Leu Ala Val Leu Phe Ser Leu Ser His Asn Phe
            340   345   350

Glu Ser Ala Asp Arg Asp Pro Thr Ala Pro Leu Lys Lys Thr Gly Glu
            355   360   365

Pro Val Asp Trp Phe Lys Thr Gln Val Glu Thr Ser Cys Thr Tyr Gly
    370   375   380

Gly Phe Leu Ser Gly Cys Phe Thr Gly Gly Leu Asn Phe Gln Val Glu
385   390   395   400

His His Leu Phe Pro Arg Met Ser Ser Ala Trp Tyr Pro Tyr Ile Ala
            405   410   415

Pro Lys Val Arg Glu Ile Cys Ala Lys His Gly Val His Tyr Ala Tyr
            420   425   430

Tyr Pro Trp Ile His Gln Asn Phe Leu Ser Thr Val Arg Tyr Met His
    435   440   445

Ala Ala Gly Thr Gly Ala Asn Trp Arg Gln Met Ala Arg Glu Asn Pro
    450   455   460

Leu Thr Gly Arg Ala
465

<210>   15
<211>   3598
<212>   DNA
<213>   artificial sequence

<220>
<221>   misc_feature
<222>   (1)..(3598)

<223>   The sequence is a plant promoter-terminator expression cassette in vector
pUC19

<400>   15
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca   60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg   120

```
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc    180

accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc    240

attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat    300

tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta cgccagggt    360

tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt cggcgcgccg agctcctcga    420

gcaaatttac acattgccac taaacgtcta aacccttgta atttgttttt gttttactat    480

gtgtgttatg tatttgattt gcgataaatt tttatatttg gtactaaatt tataacacct    540

tttatgctaa cgtttgccaa cacttagcaa tttgcaagtt gattaattga ttctaaatta    600

tttttgtctt ctaaatacat atactaatca actggaaatg taaatatttg ctaatatttc    660

tactatagga gaattaaagt gagtgaatat ggtaccacaa ggtttggaga tttaattgtt    720

gcaatgctgc atggatggca tatacaccaa acattcaata attcttgagg ataataatgg    780

taccacacaa gatttgaggt gcatgaacgt cacgtggaca aaaggtttag taattttca    840

agacaacaat gttaccacac acaagttttg aggtgcatgc atggatgccc tgtggaaagt    900

ttaaaaatat tttggaaatg atttgcatgg aagccatgtg taaaaccatg acatccactt    960

ggaggatgca ataatgaaga aaactacaaa tttacatgca actagttatg catgtagtct   1020

atataatgag gattttgcaa tactttcatt catacacact cactaagttt tacacgatta   1080

taatttcttc atagccagcc caccgcggtg ggcggccgcc tgcagtctag aaggcctcct   1140

gctttaatga gatatgcgag acgcctatga tcgcatgata tttgctttca attctgttgt   1200

gcacgttgta aaaaacctga gcatgtgtag ctcagatcct taccgccggt ttcggttcat   1260

tctaatgaat atatcacccg ttactatcgt attttttatga ataatattct ccgttcaatt   1320

tactgattgt ccgtcgacga attcgagctc ggcgcgccaa gcttggcgta atcatggtca   1380

tagctgtttc ctgtgtgaaa ttgttatccg ctcacaattc cacacaacat acgagccgga   1440

agcataaagt gtaaagcctg gggtgcctaa tgagtgagct aactcacatt aattgcgttg   1500

cgctcactgc ccgctttcca gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc   1560

caacgcgcgg ggagaggcgg tttgcgtatt gggcgctctt ccgcttcctc gctcactgac   1620

tcgctgcgct cggtcgttcg ctgcggcga gcggtatcag ctcactcaaa ggcggtaata   1680

cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa   1740

aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct ccgcccccct   1800

gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac aggactataa   1860
```

```
agataccagg cgtttcccccc tggaagctcc ctcgtgcgct ctcctgttcc gaccctgccg      1920

cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttc tcatagctca      1980

cgctgtaggt atctcagttc ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa      2040

cccccgttc agcccgaccg ctgcgcctta tccggtaact atcgtcttga gtccaacccg       2100

gtaagacacg acttatcgcc actggcagca gccactggta acaggattag cagagcgagg      2160

tatgtaggcg gtgctacaga gttcttgaag tggtggccta actacggcta cactagaagg      2220

acagtatttg gtatctgcgc tctgctgaag ccagttacct tcggaaaaag agttggtagc      2280

tcttgatccg gcaaacaaac caccgctggt agcggtggtt tttttgtttg caagcagcag      2340

attacgcgca gaaaaaaagg atctcaagaa gatcctttga tcttttctac ggggtctgac      2400

gctcagtgga acgaaaactc acgttaaggg attttggtca tgagattatc aaaaaggatc      2460

ttcacctaga tcctttttaaa ttaaaaatga agttttaaat caatctaaag tatatatgag     2520

taaacttggt ctgacagtta ccaatgctta atcagtgagg cacctatctc agcgatctgt      2580

ctatttcgtt catccatagt tgcctgactc cccgtcgtgt agataactac gatacgggag      2640

ggcttaccat ctggccccag tgctgcaatg ataccgcgag acccacgctc accggctcca      2700

gatttatcag caataaacca gccagccgga agggccgagc gcagaagtgg tcctgcaact      2760

ttatccgcct ccatccagtc tattaattgt tgccgggaag ctagagtaag tagttcgcca      2820

gttaatagtt tgcgcaacgt tgttgccatt gctacaggca tcgtggtgtc acgctcgtcg      2880

tttggtatgg cttcattcag ctccggttcc caacgatcaa ggcgagttac atgatccccc      2940

atgttgtgca aaaaagcggt tagctccttc ggtcctccga tcgttgtcag aagtaagttg      3000

gccgcagtgt tatcactcat ggttatggca gcactgcata attctcttac tgtcatgcca      3060

tccgtaagat gcttttctgt gactggtgag tactcaacca agtcattctg agaatagtgt      3120

atgcggcgac cgagttgctc ttgcccggcg tcaatacggg ataataccgc gccacatagc      3180

agaactttaa aagtgctcat cattggaaaa cgttcttcgg ggcgaaaact ctcaaggatc      3240

ttaccgctgt tgagatccag ttcgatgtaa cccactcgtg cacccaactg atcttcagca      3300

tctttactt tcaccagcgt ttctgggtga gcaaaaacag gaaggcaaaa tgccgcaaaa       3360

aagggaataa gggcgacacg gaaatgttga atactcatac tcttccttt tcaatattat       3420

tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa      3480

aataaacaaa taggggttcc gcgcacattt ccccgaaaag tgccacctga cgtctaagaa      3540
```

accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttcgtc   3598

<210> 16
<211> 3590
<212> DNA
<213> artificial sequence

<220>
<221> misc_feature
<222> (1)..(3590)

<223> The sequence is a plant promoter-terminator expression cassette in vector pUC19

<400> 16
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca   60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg   120

ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc   180

accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc   240

attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat   300

tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta cgccagggt   360

tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt cggcgcgccg agctcctcga   420

gcaaatttac acattgccac taaacgtcta aacccttgta atttgttttt gttttactat   480

gtgtgttatg tatttgattt gcgataaatt tttatatttg gtactaaatt tataacacct   540

tttatgctaa cgtttgccaa cacttagcaa tttgcaagtt gattaattga ttctaaatta   600

tttttgtctt ctaaatacat atactaatca actggaaatg taaatatttg ctaatatttc   660

tactatagga gaattaaagt gagtgaatat ggtaccacaa ggtttggaga tttaattgtt   720

gcaatgctgc atggatggca tatacaccaa acattcaata attcttgagg ataataatgg   780

taccacacaa gatttgaggt gcatgaacgt cacgtggaca aaaggtttag taatttttca   840

agacaacaat gttaccacac acaagttttg aggtgcatgc atggatgccc tgtggaaagt   900

ttaaaaatat tttggaaatg atttgcatgg aagccatgtg taaaaccatg acatccactt   960

ggaggatgca ataatgaaga aaactacaaa tttacatgca actagttatg catgtagtct   1020

atataatgag gattttgcaa tactttcatt catacacact cactaagttt tacacgatta   1080

taatttcttc atagccagcg gatccgatat cgggcccgct agcgttaacc ctgctttaat   1140

gagatatgcg agacgcctat gatcgcatga tatttgcttt caattctgtt gtgcacgttg   1200

```
taaaaaacct gagcatgtgt agctcagatc cttaccgccg gtttcggttc attctaatga   1260

atatatcacc cgttactatc gtattttat gaataatatt ctccgttcaa tttactgatt   1320

gtccgtcgac gaattcgagc tcggcgcgcc aagcttggcg taatcatggt catagctgtt   1380

tcctgtgtga aattgttatc cgctcacaat tccacacaac atacgagccg gaagcataaa   1440

gtgtaaagcc tggggtgcct aatgagtgag ctaactcaca ttaattgcgt tgcgctcact   1500

gcccgctttc cagtcgggaa acctgtcgtg ccagctgcat taatgaatcg gccaacgcgc   1560

ggggagaggc ggtttgcgta ttgggcgctc ttccgcttcc tcgctcactg actcgctgcg   1620

ctcggtcgtt cggctgcggc gagcggtatc agctcactca aaggcggtaa tacggttatc   1680

cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag   1740

gaaccgtaaa aaggccgcgt tgctggcgtt tttccatagg ctccgccccc ctgacgagca   1800

tcacaaaaat cgacgctcaa gtcagaggtg gcgaaacccg acaggactat aaagatacca   1860

ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc cgcttaccgg   1920

atacctgtcc gcctttctcc cttcgggaag cgtggcgctt tctcatagct cacgctgtag   1980

gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg aacccccgt   2040

tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc cggtaagaca   2100

cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg   2160

cggtgctaca gagttcttga agtggtggcc taactacggc tacactagaa ggacagtatt   2220

tggtatctgc gctctgctga agccagttac cttcggaaaa agagttggta gctcttgatc   2280

cggcaaacaa accaccgctg gtagcggtgg ttttttttgtt tgcaagcagc agattacgcg   2340

cagaaaaaaa ggatctcaag aagatccttt gatcttttct acggggtctg acgctcagtg   2400

gaacgaaaac tcacgttaag ggattttggt catgagatta tcaaaaagga tcttcaccta   2460

gatccttta aattaaaaat gaagttttaa atcaatctaa agtatatatg agtaaacttg   2520

gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct gtctatttcg   2580

ttcatccata gttgcctgac tccccgtcgt gtagataact acgatacggg agggcttacc   2640

atctggcccc agtgctgcaa tgataccgcg agacccacgc tcaccggctc cagatttatc   2700

agcaataaac cagccagccg gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc   2760

ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc cagttaatag   2820

tttgcgcaac gttgttgcca ttgctacagg catcgtggtg tcacgctcgt cgtttggtat   2880
```

```
ggcttcattc agctccggtt cccaacgatc aaggcgagtt acatgatccc ccatgttgtg    2940

caaaaaagcg gttagctcct tcggtcctcc gatcgttgtc agaagtaagt tggccgcagt    3000

gttatcactc atggttatgg cagcactgca taattctctt actgtcatgc catccgtaag    3060

atgcttttct gtgactggtg agtactcaac caagtcattc tgagaatagt gtatgcggcg    3120

accgagttgc tcttgcccgg cgtcaatacg ggataatacc gcgccacata gcagaacttt    3180

aaaagtgctc atcattggaa aacgttcttc ggggcgaaaa ctctcaagga tcttaccgct    3240

gttgagatcc agttcgatgt aacccactcg tgcacccaac tgatcttcag catcttttac    3300

tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa aaaagggaat    3360

aagggcgaca cggaaatgtt gaatactcat actcttcctt tttcaatatt attgaagcat    3420

ttatcagggt tattgtctca tgagcggata catatttgaa tgtatttaga aaaataaaca    3480

aataggggtt ccgcgcacat ttccccgaaa agtgccacct gacgtctaag aaaccattat    3540

tatcatgaca ttaacctata aaaataggcg tatcacgagg cccttttcgtc    3590
```

```
<210>  17
<211>  3584
<212>  DNA
<213>  artificial sequence

<220>
<221>  misc_feature
<222>  (1)..(3584)

<223>  The sequence is a plant promoter-terminator expression cassette in vector
pUC19


<400>  17
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca    60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg    120

ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc    180

accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc    240

attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat    300

tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta cgccagggt    360

ttccccagtc acgacgttgt aaaacgacgg ccagtgaatt cggcgcgccg agctcctcga    420

gcaaatttac acattgccac taaacgtcta aacccttgta atttgttttt gttttactat    480

gtgtgttatg tatttgattt gcgataaatt tttatatttg gtactaaatt tataacacct    540
```

```
tttatgctaa cgtttgccaa cacttagcaa tttgcaagtt gattaattga ttctaaatta    600

tttttgtctt ctaaatacat atactaatca actggaaatg taaatatttg ctaatatttc    660

tactatagga gaattaaagt gagtgaatat ggtaccacaa ggtttggaga tttaattgtt    720

gcaatgctgc atggatggca tatacaccaa acattcaata attcttgagg ataataatgg    780

taccacacaa gatttgaggt gcatgaacgt cacgtggaca aaaggtttag taattttttca    840

agacaacaat gttaccacac acaagttttg aggtgcatgc atggatgccc tgtggaaagt    900

ttaaaaatat tttggaaatg atttgcatgg aagccatgtg taaaaccatg acatccactt    960

ggaggatgca ataatgaaga aaactacaaa tttacatgca actagttatg catgtagtct   1020

atataatgag gattttgcaa tactttcatt catacacact cactaagttt tacacgatta   1080

taatttcttc atagccagca gatctgccgg catcgatccc gggccatggc ctgctttaat   1140

gagatatgcg agacgcctat gatcgcatga tatttgcttt caattctgtt gtgcacgttg   1200

taaaaaacct gagcatgtgt agctcagatc cttaccgccg gtttcggttc attctaatga   1260

atatatcacc cgttactatc gtatttttat gaataatatt ctccgttcaa tttactgatt   1320

gtccgtcgac gagctcggcg cgccaagctt ggcgtaatca tggtcatagc tgtttcctgt   1380

gtgaaattgt tatccgctca caattccaca caacatacga gccggaagca taaagtgtaa   1440

agcctggggt gcctaatgag tgagctaact cacattaatt gcgttgcgct cactgcccgc   1500

tttccagtcg ggaaacctgt cgtgccagct gcattaatga atcggccaac gcgcggggag   1560

aggcggtttg cgtattgggc gctcttccgc ttcctcgctc actgactcgc tgcgctcggt   1620

cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt tatccacaga   1680

atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg   1740

taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccccctgacg agcatcacaa   1800

aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt   1860

tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct   1920

gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct   1980

cagttcggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc   2040

cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt   2100

atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc   2160

tacagagttc ttgaagtggt ggcctaacta cggctacact agaaggacag tatttggtat   2220

ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa   2280
```

```
acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta.cgcgcagaaa    2340

aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga    2400

aaactcacgt taagggattt tggtcatgag attatcaaaa aggatcttca cctagatcct    2460

tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa cttggtctga    2520

cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcatc    2580

catagttgcc tgactccccg tcgtgtagat aactacgata cgggagggct taccatctgg    2640

ccccagtgct gcaatgatac cgcgagaccc acgctcaccg gctccagatt tatcagcaat    2700

aaaccagcca gccggaaggg ccgagcgcag aagtggtcct gcaactttat ccgcctccat    2760

ccagtctatt aattgttgcc gggaagctag agtaagtagt tcgccagtta atagtttgcg    2820

caacgttgtt.gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg gtatggcttc    2880

attcagctcc ggttcccaac gatcaaggcg agttacatga tcccccatgt tgtgcaaaaa    2940

agcggttagc tccttcggtc ctccgatcgt tgtcagaagt aagttggccg cagtgttatc    3000

actcatggtt atggcagcac tgcataattc tcttactgtc atgccatccg taagatgctt    3060

ttctgtgact ggtgagtact caaccaagtc attctgagaa tagtgtatgc ggcgaccgag    3120

ttgctcttgc ccggcgtcaa tacgggataa taccgcgcca catagcagaa ctttaaaagt    3180

gctcatcatt ggaaaacgtt cttcggggcg aaaactctca aggatcttac cgctgttgag    3240

atccagttcg atgtaaccca ctcgtgcacc caactgatct tcagcatctt ttactttcac    3300

cagcgtttct gggtgagcaa aaacaggaag gcaaaatgcc gcaaaaaagg gaataagggc    3360

gacacggaaa tgttgaatac tcatactctt cctttttcaa tattattgaa gcatttatca    3420

gggttattgt ctcatgagcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg    3480

ggttccgcgc acatttcccc gaaaagtgcc acctgacgtc taagaaacca ttattatcat    3540

gacattaacc tataaaaata ggcgtatcac gaggcccttt cgtc                     3584
```

```
<210>  18
<211>  4507
<212>  DNA
<213>  artificial sequence

<220>
<221>  misc_feature
<222>  (1)..(4507)

<223>  The sequence is a plant promoter-terminator expression cassette in vector
pUC19
```

<400> 18

```
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca       60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg      120

ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc      180

accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc      240

attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat      300

tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta cgccagggt      360

tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt cggcgcgccg agctcctcga      420

gcaaatttac acattgccac taaacgtcta aacccttgta atttgttttt gttttactat      480

gtgtgttatg tatttgattt gcgataaatt tttatatttg gtactaaatt tataacacct      540

tttatgctaa cgtttgccaa cacttagcaa tttgcaagtt gattaattga ttctaaatta      600

tttttgtctt ctaaatacat atactaatca actggaaatg taaatatttg ctaatatttc      660

tactatagga gaattaaagt gagtgaatat ggtaccacaa ggtttggaga tttaattgtt      720

gcaatgctgc atggatggca tatacaccaa acattcaata attcttgagg ataataatgg      780

taccacacaa gatttgaggt gcatgaacgt cacgtggaca aaaggtttag taatttttca      840

agacaacaat gttaccacac acaagttttg aggtgcatgc atggatgccc tgtggaaagt      900

ttaaaaatat tttggaaatg atttgcatgg aagccatgtg taaaaccatg acatccactt      960

ggaggatgca ataatgaaga aaactacaaa tttacatgca actagttatg catgtagtct     1020

atataatgag gattttgcaa tactttcatt catacacact cactaagttt tacacgatta     1080

taatttcttc atagccagcc caccgcggtg ggcggccgcc tgcagtctag aaggcctcct     1140

gctttaatga gatatgcgag acgcctatga tcgcatgata tttgctttca attctgttgt     1200

gcacgttgta aaaaacctga gcatgtgtag ctcagatcct taccgccggt ttcggttcat     1260

tctaatgaat atatcacccg ttactatcgt attttatga ataatattct ccgttcaatt     1320

tactgattgt ccgtcgagca aatttacaca ttgccactaa acgtctaaac ccttgtaatt     1380

tgttttgtt ttactatgtg tgttatgtat ttgatttgcg ataaattttt atatttggta     1440

ctaaatttat aacacctttt atgctaacgt ttgccaacac ttagcaattt gcaagttgat     1500

taattgattc taaattattt ttgtcttcta aatacatata ctaatcaact ggaaatgtaa     1560
```

```
atatttgcta atatttctac tataggagaa ttaaagtgag tgaatatggt accacaaggt    1620

ttggagattt aattgttgca atgctgcatg gatggcatat acaccaaaca ttcaataatt    1680

cttgaggata ataatggtac cacacaagat ttgaggtgca tgaacgtcac gtggacaaaa    1740

ggtttagtaa tttttcaaga caacaatgtt accacacaca agttttgagg tgcatgcatg    1800

gatgccctgt ggaaagttta aaaatatttt ggaaatgatt tgcatggaag ccatgtgtaa    1860

aaccatgaca tccacttgga ggatgcaata atgaagaaaa ctacaaattt acatgcaact    1920

agttatgcat gtagtctata taatgaggat tttgcaatac tttcattcat acacactcac    1980

taagttttac acgattataa tttcttcata gccagcggat ccgatatcgg gcccgctagc    2040

gttaaccctg ctttaatgag atatgcgaga cgcctatgat cgcatgatat ttgctttcaa    2100

ttctgttgtg cacgttgtaa aaaacctgag catgtgtagc tcagatcctt accgccggtt    2160

tcggttcatt ctaatgaata tatcacccgt tactatcgta tttttatgaa taatattctc    2220

cgttcaattt actgattgtc cgtcgacgaa ttcgagctcg gcgcgccaag cttggcgtaa    2280

tcatggtcat agctgtttcc tgtgtgaaat tgttatccgc tcacaattcc acacaacata    2340

cgagccggaa gcataaagtg taaagcctgg ggtgcctaat gagtgagcta actcacatta    2400

attgcgttgc gctcactgcc cgctttccag tcgggaaacc tgtcgtgcca gctgcattaa    2460

tgaatcggcc aacgcgcggg gagaggcggt ttgcgtattg ggcgctcttc cgcttcctcg    2520

ctcactgact cgctgcgctc ggtcgttcgg ctgcggcgag cggtatcagc tcactcaaag    2580

gcggtaatac ggttatccac agaatcaggg gataacgcag gaaagaacat gtgagcaaaa    2640

ggccagcaaa aggccaggaa ccgtaaaaag ccgcgttgc tggcgttttt ccataggctc    2700

cgcccccctg acgagcatca caaaaatcga cgctcaagtc agaggtggcg aaacccgaca    2760

ggactataaa gataccaggc gtttccccct ggaagctccc tcgtgcgctc tcctgttccg    2820

accctgccgc ttaccggata cctgtccgcc tttctccctt cgggaagcgt ggcgctttct    2880

catagctcac gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt    2940

gtgcacgaac ccccgttca gcccgaccgc tgcgccttat ccggtaacta tcgtcttgag    3000

tccaacccgg taagacacga cttatcgcca ctggcagcag ccactggtaa caggattagc    3060

agagcgaggt atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa ctacggctac    3120

actagaagga cagtatttgg tatctgcgct ctgctgaagc cagttacctt cggaaaaaga    3180

gttggtagct cttgatccgg caaacaaacc accgctggta gcggtggttt ttttgtttgc    3240

aagcagcaga ttacgcgcag aaaaaaagga tctcaagaag atcctttgat cttttctacg    3300
```

```
gggtctgacg ctcagtggaa cgaaaactca cgttaaggga ttttggtcat gagattatca    3360

aaaaggatct tcacctagat ccttttaaat taaaaatgaa gttttaaatc aatctaaagt    3420

atatatgagt aaacttggtc tgacagttac caatgcttaa tcagtgaggc acctatctca    3480

gcgatctgtc tatttcgttc atccatagtt gcctgactcc ccgtcgtgta gataactacg    3540

atacgggagg gcttaccatc tggccccagt gctgcaatga taccgcgaga cccacgctca    3600

ccggctccag atttatcagc aataaaccag ccagccggaa gggccgagcg cagaagtggt    3660

cctgcaactt tatccgcctc catccagtct attaattgtt gccgggaagc tagagtaagt    3720

agttcgccag ttaatagttt gcgcaacgtt gttgccattg ctacaggcat cgtggtgtca    3780

cgctcgtcgt ttggtatggc ttcattcagc tccggttccc aacgatcaag gcgagttaca    3840

tgatccccca tgttgtgcaa aaaagcggtt agctccttcg gtcctccgat cgttgtcaga    3900

agtaagttgg ccgcagtgtt atcactcatg gttatggcag cactgcataa ttctcttact    3960

gtcatgccat ccgtaagatg cttttctgtg actggtgagt actcaaccaa gtcattctga    4020

gaatagtgta tgcggcgacc gagttgctct tgcccggcgt caatacggga taataccgcg    4080

ccacatagca gaactttaaa agtgctcatc attggaaaac gttcttcggg gcgaaaactc    4140

tcaaggatct taccgctgtt gagatccagt tcgatgtaac ccactcgtgc acccaactga    4200

tcttcagcat cttttacttt caccagcgtt tctgggtgag caaaaacagg aaggcaaaat    4260

gccgcaaaaa agggaataag ggcgacacgg aaatgttgaa tactcatact cttccttttt    4320

caatattatt gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt    4380

atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt gccacctgac    4440

gtctaagaaa ccattattat catgacatta acctataaaa ataggcgtat cacgaggccc    4500

tttcgtc                                                              4507


<210>   19
<211>   17752
<212>   DNA
<213>   Phaeodactylum tricornutum, Physcomitrella patens

<220>
<221>   CDS
<222>   (11543)..(12415)
<223>   Delta-6-elongase


<220>
<221>   CDS
```

```
<222>  (13313)..(14890)
<223>  Delta-6-desaturase


<220>
<221>  CDS
<222>  (15791)..(17200)
<223>  Delta-5-desaturase


<400>  19
gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg atccgacagc      60

gcgcccagca caggtgcgca ggcaaattgc accaacgcat acagcgccag cagaatgcca     120

tagtgggcgg tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccgg cagcaccggc     180

ataatcaggc cgatgccgac agcgtcgagc gcgacagtgc tcagaattac gatcaggggt     240

atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg aacgcgcgga     300

ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa gtgtcaagca     360

tgacaaagtt gcagccgaat acagtgatcc gtgccgccct ggacctgttg aacgaggtcg     420

gcgtagacgg tctgacgaca cgcaaactgg cggaacggtt ggggggttcag cagccggcgc     480

tttactggca cttcaggaac aagcgggcgc tgctcgacgc actggccgaa gccatgctgg     540

cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca tttctgatcg     600

ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg cgcatccatg     660

ccggcacgcg accgggcgca ccgcagatgg aaacggccga cgcgcagctt cgcttcctct     720

gcgaggcggg ttttcggcc ggggacgccg tcaatgcgct gatgacaatc agctacttca     780

ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag cgcggcggca     840

ccgttgaaca ggctccgctc tcgccgctgt tgcgggccgc gatagacgcc ttcgacgaag     900

ccggtccgga cgcagcgttc gagcagggac tcgcggtgat tgtcgatgga ttggcgaaaa     960

ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc    1020

tgccggagcg caacccactc actacagcag agccatgtag acaacatccc ctcccccttt    1080

ccaccgcgtc agacgcccgt agcagcccgc tacgggcttt ttcatgccct gccctagcgt    1140

ccaagcctca cggccgcgct cggcctctct ggcggccttc tggcgctctt ccgcttcctc    1200

gctcactgac tcgctgcgct cggtcgttcg gctgcggcga gcggtatcag ctcactcaaa    1260

ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa    1320

aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct    1380
```

```
ccgccccct gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac   1440

aggactataa agataccagg cgtttccccc tggaagctcc ctcgtgcgct ctcctgttcc   1500

gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttt   1560

ccgctgcata accctgcttc ggggtcatta tagcgatttt ttcggtatat ccatcctttt   1620

tcgcacgata tacaggattt tgccaaaggg ttcgtgtaga ctttccttgg tgtatccaac   1680

ggcgtcagcc gggcaggata ggtgaagtag gcccacccgc gagcgggtgt tccttcttca   1740

ctgtccctta ttcgcacctg gcggtgctca acgggaatcc tgctctgcga ggctggccgg   1800

ctaccgccgg cgtaacagat gagggcaagc ggatggctga tgaaaccaag ccaaccagga   1860

agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa   1920

aggcggcggc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc caggctaca   1980

aaatcacggg cgtcgtggac tatgagcacg tccgcgagct ggcccgcatc aatggcgacc   2040

tgggccgcct gggcggcctg ctgaaactct ggctcaccga cgacccgcgc acggcgcggt   2100

tcggtgatgc cacgatcctc gccctgctgg cgaagatcga agagaagcag gacgagcttg   2160

gcaaggtcat gatgggcgtg gtccgcccga gggcagagcc atgacttttt tagccgctaa   2220

aacggccggg gggtgcgcgt gattgccaag cacgtcccca tgcgctccat caagaagagc   2280

gacttcgcgg agctggtgaa gtacatcacc gacgagcaag gcaagaccga gcgcctttgc   2340

gacgctcacc gggctggttg ccctcgccgc tgggctggcg gccgtctatg gccctgcaaa   2400

cgcgccagaa acgccgtcga agccgtgtgc gagacaccgc ggccgccggc gttgtggata   2460

cctcgcggaa aacttggccc tcactgacag atgaggggcg gacgttgaca cttgaggggc   2520

cgactcaccc ggcgcggcgt tgacagatga ggggcaggct cgatttcggc cggcgacgtg   2580

gagctggcca gcctcgcaaa tcggcgaaaa cgcctgattt tacgcgagtt tcccacagat   2640

gatgtggaca agcctgggga taagtgccct gcggtattga cacttgaggg gcgcgactac   2700

tgacagatga ggggcgcgat ccttgacact tgaggggcag agtgctgaca gatgaggggc   2760

gcacctattg acatttgagg ggctgtccac aggcagaaaa tccagcattt gcaagggttt   2820

ccgcccgttt ttcggccacc gctaacctgt cttttaacct gcttttaaac caatatttat   2880

aaaccttgtt tttaaccagg gctgcgccct gtgcgcgtga ccgcgcacgc cgaaggggggg  2940

tgcccccct tctcgaaccc tcccggcccg ctaacgcggg cctcccatcc ccccagggg c  3000

tgcgcccctc ggccgcgaac ggcctcaccc caaaaatggc agcgctggca gtccttgcca   3060

ttgccgggat cggggcagta acgggatggg cgatcagccc gagcgcgacg cccggaagca   3120
```

```
ttgacgtgcc gcaggtgctg gcatcgacat tcagcgacca ggtgccgggc agtgagggcg      3180

gcggcctggg tggcggcctg cccttcactt cggccgtcgg ggcattcacg gacttcatgg      3240

cggggccggc aattttacc ttgggcattc ttggcatagt ggtcgcgggt gccgtgctcg       3300

tgttcggggg tgcgataaac ccagcgaacc atttgaggtg ataggtaaga ttataccgag      3360

gtatgaaaac gagaattgga cctttacaga attactctat gaagcgccat atttaaaaag      3420

ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgacaa      3480

tactgataag ataatatatc ttttatatag aagatatcgc cgtatgtaag gatttcaggg      3540

ggcaaggcat aggcagcgcg cttatcaata tatctataga atgggcaaag cataaaaact      3600

tgcatggact aatgcttgaa acccaggaca ataaccttat agcttgtaaa ttctatcata      3660

attgggtaat gactccaact tattgatagt gttttatgtt cagataatgc ccgatgactt      3720

tgtcatgcag ctccaccgat tttgagaacg acagcgactt ccgtcccagc cgtgccaggt      3780

gctgcctcag attcaggtta tgccgctcaa ttcgctgcgt atatcgcttg ctgattacgt      3840

gcagctttcc cttcaggcgg gattcataca gcggccagcc atccgtcatc catatcacca      3900

cgtcaaaggg tgacagcagg ctcataagac gccccagcgt cgccatagtg cgttcaccga      3960

atacgtgcgc aacaaccgtc ttccggagac tgtcatacgc gtaaaacagc cagcgctggc      4020

gcgatttagc cccgacatag ccccactgtt cgtccatttc cgcgcagacg atgacgtcac      4080

tgcccggctg tatgcgcgag gttaccgact gcggcctgag ttttttaagt gacgtaaaat      4140

cgtgttgagg ccaacgccca taatgcgggc tgttgcccgg catccaacgc cattcatggc      4200

catatcaatg attttctggt gcgtaccggg ttgagaagcg gtgtaagtga actgcagttg      4260

ccatgtttta cggcagtgag agcagagata gcgctgatgt ccggcggtgc ttttgccgtt      4320

acgcaccacc ccgtcagtag ctgaacagga gggacagctg atagacacag aagccactgg      4380

agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca cccataattg      4440

tggtttcaaa atcggctccg tcgatactat gttatacgcc aactttgaaa acaactttga      4500

aaaagctgtt ttctggtatt taaggtttta gaatgcaagg aacagtgaat tggagttcgt      4560

cttgttataa ttagcttctt ggggtatctt taaatactgt agaaagagg aaggaaataa       4620

taaatggcta aaatgagaat atcaccggaa ttgaaaaaac tgatcgaaaa ataccgctgc      4680

gtaaaagata cggaaggaat gtctcctgct aaggtatata agctggtggg agaaaatgaa      4740

aacctatatt taaaaatgac ggacagccgg tataaaggga ccacctatga tgtggaacgg      4800
```

```
gaaaaggaca tgatgctatg gctggaagga aagctgcctg ttccaaaggt cctgcacttt    4860

gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt cctttgctcg    4920

gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcatc    4980

aggctctttc actccatcga catatcggat tgtccctata cgaatagctt agacagccgc    5040

ttagccgaat tggattactt actgaataac gatctggccg atgtggattg cgaaaactgg    5100

gaagaagaca ctccatttaa agatccgcgc gagctgtatg attttttaaa gacggaaaag    5160

cccgaagagg aacttgtctt ttcccacggc gacctgggag acagcaacat ctttgtgaaa    5220

gatggcaaag taagtggctt tattgatctt gggagaagcg gcagggcgga caagtggtat    5280

gacattgcct tctgcgtccg gtcgatcagg gaggatatcg gggaagaaca gtatgtcgag    5340

ctattttttg acttactggg gatcaagcct gattgggaga aaataaaata ttatatttta    5400

ctggatgaat tgttttagta cctagatgtg gcgcaacgat gccggcgaca agcaggagcg    5460

caccgacttc ttccgcatca agtgttttgg ctctcaggcc gaggcccacg gcaagtattt    5520

gggcaagggg tcgctggtat tcgtgcaggg caagattcgg aataccaagt acgagaagga    5580

cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc tggacaccaa    5640

ggcaccaggc gggtcaaatc aggaataagg gcacattgcc ccggcgtgag tcggggcaat    5700

cccgcaagga gggtgaatga atcggacgtt tgaccggaag gcatacaggc aagaactgat    5760

cgacgcgggg ttttccgccg aggatgccga aaccatcgca agccgcaccg tcatgcgtgc    5820

gccccgcgaa accttccagt ccgtcggctc gatggtccag caagctacgg ccaagatcga    5880

gcgcgacagc gtgcaactgg ctccccctgc cctgcccgcg ccatcggccg ccgtggagcg    5940

ttcgcgtcgt ctcgaacagg aggcggcagg tttggcgaag tcgatgacca tcgacacgcg    6000

aggaactatg acgaccaaga agcgaaaaac cgccggcgag gacctggcaa aacaggtcag    6060

cgaggccaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg aaatgcagct    6120

ttccttgttc gatattgcgc cgtggccgga cacgatgcga gcgatgccaa cgacacggc     6180

ccgctctgcc ctgttcacca cgcgcaacaa gaaaatcccg cgcgaggcgc tgcaaaacaa    6240

ggtcattttc cacgtcaaca aggacgtgaa gatcacctac accggcgtcg agctgcgggc    6300

cgacgatgac gaactggtgt ggcagcaggt gttggagtac gcgaagcgca cccctatcgg    6360

cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggt cgatcaatgg    6420

ccggtattac acgaaggccg aggaatgcct gtcgcgccta caggcgacgg cgatgggctt    6480

cacgtccgac cgcgttgggc acctggaatc ggtgtcgctg ctgcaccgct tccgcgtcct    6540
```

```
ggaccgtggc aagaaaacgt cccgttgcca ggtcctgatc gacgaggaaa tcgtcgtgct    6600

gtttgctggc gaccactaca cgaaattcat atgggagaag taccgcaagc tgtcgccgac    6660

ggcccgacgg atgttcgact atttcagctc gcaccgggag ccgtacccgc tcaagctgga    6720

aaccttccgc ctcatgtgcg gatcggattc cacccgcgtg aagaagtggc gcgagcaggt    6780

cggcgaagcc tgcgaagagt tgcgaggcag cggcctggtg gaacacgcct gggtcaatga    6840

tgacctggtg cattgcaaac gctagggcct tgtggggtca gttccggctg ggggttcagc    6900

agccagcgct ttactggcat ttcaggaaca agcgggcact gctcgacgca cttgcttcgc    6960

tcagtatcgc tcgggacgca cggcgcgctc tacgaactgc cgataaacag aggattaaaa    7020

ttgacaattg tgattaaggc tcagattcga cggcttggag cggccgacgt gcaggatttc    7080

cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tgttcgggtc cgtttacgag    7140

cacgaggaga aaaagcccat ggaggcgttc gctgaacggt tgcgagatgc cgtggcattc    7200

ggcgcctaca tcgacggcga gatcattggg ctgtcggtct tcaaacagga ggacggcccc    7260

aaggacgctc acaaggcgca tctgtccggc gttttcgtgg agcccgaaca gcgaggccga    7320

ggggtcgccg gtatgctgct gcgggcgttg ccggcgggtt tattgctcgt gatgatcgtc    7380

cgacagattc caacgggaat ctggtggatg cgcatcttca tcctcggcgc acttaatatt    7440

tcgctattct ggagcttgtt gtttatttcg gtctaccgcc tgccgggcgg ggtcgcggcg    7500

acggtaggcg ctgtgcagcc gctgatggtc gtgttcatct ctgccgctct gctaggtagc    7560

ccgatacgat tgatggcggt cctgggggct atttgcggaa ctgcgggcgt ggcgctgttg    7620

gtgttgacac caaacgcagc gctagatcct gtcggcgtcg cagcgggcct ggcggggggcg    7680

gtttccatgg cgttcggaac cgtgctgacc cgcaagtggc aacctcccgt gcctctgctc    7740

acctttaccg cctggcaact ggcggccgga ggacttctgc tcgttccagt agctttagtg    7800

tttgatccgc caatcccgat gcctacagga accaatgttc tcggcctggc gtggctcggc    7860

ctgatcggag cgggtttaac ctacttcctt tggttccggg ggatctcgcg actcgaacct    7920

acagttgttt ccttactggg ctttctcagc cccagatctg gggtcgatca gccggggatg    7980

catcaggccg acagtcggaa cttcgggtcc ccgacctgta ccattcggtg agcaatggat    8040

aggggagttg atatcgtcaa cgttcacttc taaagaaata gcgccactca gcttcctcag    8100

cggctttatc cagcgatttc ctattatgtc ggcatagttc tcaagatcga cagcctgtca    8160

cggttaagcg agaaatgaat aagaaggctg ataattcgga tctctgcgag ggagatgata    8220
```

```
tttgatcaca ggcagcaacg ctctgtcatc gttacaatca acatgctacc ctccgcgaga    8280

tcatccgtgt ttcaaacccg gcagcttagt tgccgttctt ccgaatagca tcggtaacat    8340

gagcaaagtc tgccgcctta caacggctct cccgctgacg ccgtcccgga ctgatgggct    8400

gcctgtatcg agtggtgatt ttgtgccgag ctgccggtcg gggagctgtt ggctggctgg    8460

tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg    8520

gacgttttta atgtactggg gtggtttttc ttttcaccag tgagacgggc aacagctgat    8580

tgcccttcac cgcctggccc tgagagagtt gcagcaagcg tccacgctg gtttgcccca    8640

gcaggcgaaa atcctgtttg atggtggttc cgaaatcggc aaaatccctt ataaatcaaa    8700

agaatagccc gagatagggt tgagtgttgt tccagtttgg aacaagagtc cactattaaa    8760

gaacgtggac tccaacgtca aagggcgaaa aaccgtctat cagggcgatg gcccactacg    8820

tgaaccatca cccaaatcaa gttttttggg gtcgaggtgc cgtaaagcac taaatcggaa    8880

ccctaaaggg agcccccgat ttagagcttg acggggaaag ccggcgaacg tggcgagaaa    8940

ggaagggaag aaagcgaaag gagcgggcgc cattcaggct gcgcaactgt tgggaagggc    9000

gatcggtgcg ggcctcttcg ctattacgcc agctggcgaa aggggatgt gctgcaaggc    9060

gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg    9120

aattaattcc catcttgaaa gaaatatagt ttaaatattt attgataaaa taacaagtca    9180

ggtattatag tccaagcaaa aacataaatt tattgatgca agtttaaatt cagaaatatt    9240

tcaataactg attatatcag ctggtacatt gccgtagatg aaagactgag tgcgatatta    9300

tgtgtaatac ataaattgat gatatagcta gcttagctca tcgggggatc cgtcgaagct    9360

agcttgggtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat gcgctgcgaa    9420

tcgggagcgg cgataccgta aagcacgagg aagcggtcag cccattcgcc gccaagctct    9480

tcagcaatat cacgggtagc caacgctatg tcctgatagc ggtccgccac acccagccgg    9540

ccacagtcga tgaatccaga aaagcggcca ttttccacca tgatattcgg caagcaggca    9600

tcgccatggg tcacgacgag atcctcgccg tcgggcatgc gcgccttgag cctggcgaac    9660

agttcggctg gcgcgagccc ctgatgctct tcgtccagat catcctgatc gacaagaccg    9720

gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg cttggtggtc gaatgggcag    9780

gtagccggat caagcgtatg cagccgccgc attgcatcag ccatgatgga tactttctcg    9840

gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgcccaa tagcagccag    9900

tcccttcccg cttcagtgac aacgtcgagc acagctgcgc aaggaacgcc cgtcgtggcc    9960
```

```
agccacgata gccgcgctgc ctcgtcctgc agttcattca gggcaccgga caggtcggtc   10020

ttgacaaaaa gaaccgggcg cccctgcgct gacagccgga acacggcggc atcagagcag   10080

ccgattgtct gttgtgccca gtcatagccg aatagcctct ccacccaagc ggccggagaa   10140

cctgcgtgca atccatcttg ttcaatccaa gctcccatgg ccctcgact agagtcgaga    10200

tctggattga gagtgaatat gagactctaa ttggataccg aggggaattt atggaacgtc   10260

agtggagcat ttttgacaag aaatatttgc tagctgatag tgaccttagg cgacttttga   10320

acgcgcaata atggtttctg acgtatgtgc ttagctcatt aaactccaga aacccgcggc   10380

tgagtggctc cttcaacgtt gcggttctgt cagttccaaa cgtaaaacgg cttgtcccgc   10440

gtcatcggcg ggggtcataa cgtgactccc ttaattctcc gctcatgatc ttgatcccct   10500

gcgccatcag atccttggcg gcaagaaagc catccagttt actttgcagg gcttcccaac   10560

cttaccagag ggcgccccag ctggcaattc cggttcgctt gctgtccata aaaccgccca   10620

gtctagctat cgccatgtaa gcccactgca agctacctgc tttctctttg cgcttgcgtt   10680

ttcccttgtc cagatagccc agtagctgac attcatccgg ggtcagcacc gtttctgcgg   10740

actggctttc tacgtgttcc gcttccttta gcagcccttg cgccctgagt gcttgcggca   10800

gcgtgaagct tgcatgcctg caggtcgacg gcgcgccgag ctcctcgagc aaatttacac   10860

attgccacta aacgtctaaa cccttgtaat ttgttttgt tttactatgt gtgttatgta    10920

tttgatttgc gataaatttt tatatttggt actaaattta taacaccttt tatgctaacg   10980

tttgccaaca cttagcaatt tgcaagttga ttaattgatt ctaaattatt tttgtcttct   11040

aaatacatat actaatcaac tggaaatgta aatatttgct aatatttcta ctataggaga   11100

attaaagtga gtgaatatgg taccacaagg tttggagatt taattgttgc aatgctgcat   11160

ggatggcata tacaccaaac attcaataat tcttgaggat aataatggta ccacacaaga   11220

tttgaggtgc atgaacgtca cgtggacaaa aggtttagta attttttcaag acaacaatgt  11280

taccacacac aagtttttgag gtgcatgcat ggatgccctg tggaaagttt aaaaatattt   11340

tggaaatgat ttgcatggaa gccatgtgta aaaccatgac atccacttgg aggatgcaat   11400

aatgaagaaa actacaaatt tacatgcaac tagttatgca tgtagtctat ataatgagga   11460

ttttgcaata ctttcattca tacacactca ctaagtttta cacgattata atttcttcat   11520

agccagccca ccgcggtgga aa atg gag gtc gtg gag aga ttc tac ggt gag   11572
                          Met Glu Val Val Glu Arg Phe Tyr Gly Glu
                           1           5                      10
```

```
ttg gat ggg aag gtc tcg cag ggc gtg aat gca ttg ctg ggt agt ttt          11620
Leu Asp Gly Lys Val Ser Gln Gly Val Asn Ala Leu Leu Gly Ser Phe
            15                  20                  25

ggg gtg gag ttg acg gat acg ccc act acc aaa ggc ttg ccc ctc gtt          11668
Gly Val Glu Leu Thr Asp Thr Pro Thr Thr Lys Gly Leu Pro Leu Val
            30                  35                  40

gac agt ccc aca ccc atc gtc ctc ggt gtt tct gta tac ttg act att          11716
Asp Ser Pro Thr Pro Ile Val Leu Gly Val Ser Val Tyr Leu Thr Ile
            45                  50                  55

gtc att gga ggg ctt ttg tgg ata aag gcc agg gat ctg aaa ccg cgc          11764
Val Ile Gly Gly Leu Leu Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg
        60                  65                  70

gcc tcg gag cca ttt ttg ctc caa gct ttg gtg ctt gtg cac aac ctg          11812
Ala Ser Glu Pro Phe Leu Leu Gln Ala Leu Val Leu Val His Asn Leu
75                  80                  85                  90

ttc tgt ttt gcg ctc agt ctg tat atg tgc gtg ggc atc gct tat cag          11860
Phe Cys Phe Ala Leu Ser Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln
                95                  100                 105

gct att acc tgg cgg tac tct ctc tgg ggc aat gca tac aat cct aaa          11908
Ala Ile Thr Trp Arg Tyr Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys
                110                 115                 120

cat aaa gag atg gcg att ctg gta tac ttg ttc tac atg tct aag tac          11956
His Lys Glu Met Ala Ile Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr
            125                 130                 135

gtg gaa ttc atg gat acc gtt atc atg ata ctg aag cgc agc acc agg          12004
Val Glu Phe Met Asp Thr Val Ile Met Ile Leu Lys Arg Ser Thr Arg
            140                 145                 150

caa ata agc ttc ctc cac gtt tat cat cat tct tca att tcc ctc att          12052
Gln Ile Ser Phe Leu His Val Tyr His His Ser Ser Ile Ser Leu Ile
155                 160                 165                 170

tgg tgg gct att gct cat cac gct cct ggc ggt gaa gca tat tgg tct          12100
Trp Trp Ala Ile Ala His His Ala Pro Gly Gly Glu Ala Tyr Trp Ser
                175                 180                 185

gcg gct ctg aac tca gga gtg cat gtt ctc atg tat gcg tat tac ttc          12148
Ala Ala Leu Asn Ser Gly Val His Val Leu Met Tyr Ala Tyr Tyr Phe
                190                 195                 200

ttg gct gcc tgc ctt cga agt agc cca aag tta aaa aat aag tac ctt          12196
Leu Ala Ala Cys Leu Arg Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu
            205                 210                 215

ttt tgg ggc agg tac ttg aca caa ttc caa atg ttc cag ttt atg ctg          12244
Phe Trp Gly Arg Tyr Leu Thr Gln Phe Gln Met Phe Gln Phe Met Leu
            220                 225                 230

aac tta gtg cag gct tac tac gac atg aaa acg aat gcg cca tat cca          12292
```

```
Asn Leu Val Gln Ala Tyr Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro
235             240             245             250

caa tgg ctg atc aag att ttg ttc tac tac atg atc tcg ttg ctg ttt    12340
Gln Trp Leu Ile Lys Ile Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe
                255             260             265

ctt ttc ggc aat ttt tac gta caa aaa tac atc aaa ccc tct gac gga    12388
Leu Phe Gly Asn Phe Tyr Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly
                270             275             280

aag caa aag gga gct aaa act gag tga tctagaaggc ctcctgcttt          12435
Lys Gln Lys Gly Ala Lys Thr Glu
            285             290

aatgagatat gcgagacgcc tatgatcgca tgatatttgc tttcaattct gttgtgcacg  12495

ttgtaaaaaa cctgagcatg tgtagctcag atccttaccg ccggtttcgg ttcattctaa  12555

tgaatatatc acccgttact atcgtatttt tatgaataat attctccgtt caatttactg  12615

attgtccgtc gagcaaattt acacattgcc actaaacgtc taaacccttg taatttgttt  12675

ttgttttact atgtgtgtta tgtatttgat ttgcgataaa tttttatatt tggtactaaa  12735

tttataacac cttttatgct aacgtttgcc aacacttagc aatttgcaag ttgattaatt  12795

gattctaaat tatttttgtc ttctaaatac atatactaat caactggaaa tgtaaatatt  12855

tgctaatatt tctactatag gagaattaaa gtgagtgaat atggtaccac aaggtttgga  12915

gatttaattg ttgcaatgct gcatggatgg catatacacc aaacattcaa taattcttga  12975

ggataataat ggtaccacac aagatttgag gtgcatgaac gtcacgtgga caaaaggttt  13035

agtaattttt caagacaaca atgttaccac acacaagttt tgaggtgcat gcatggatgc  13095

cctgtggaaa gtttaaaaat attttggaaa tgatttgcat ggaagccatg tgtaaaacca  13155

tgacatccac ttggaggatg caataatgaa gaaaactaca aatttacatg caactagtta  13215

tgcatgtagt ctatataatg aggattttgc aatactttca ttcatacaca ctcactaagt  13275

tttacacgat tataatttct tcatagccag cggatcc atg gta ttc gcg ggc ggt  13330
                                         Met Val Phe Ala Gly Gly
                                                             295

gga ctt cag cag ggc tct ctc gaa gaa aac atc gac gtc gag cac att    13378
Gly Leu Gln Gln Gly Ser Leu Glu Glu Asn Ile Asp Val Glu His Ile
                300             305             310

gcc agt atg tct ctc ttc agc gac ttc ttc agt tat gtg tct tca act    13426
Ala Ser Met Ser Leu Phe Ser Asp Phe Phe Ser Tyr Val Ser Ser Thr
                315             320             325

gtt ggt tcg tgg agc gta cac agt ata caa cct ttg aag cgc ctg acg    13474
Val Gly Ser Trp Ser Val His Ser Ile Gln Pro Leu Lys Arg Leu Thr
```

```
              330                    335                    340

agt aag aag cgt gtt tcg gaa agc gct gcc gtg caa tgt ata tca gct    13522
Ser Lys Lys Arg Val Ser Glu Ser Ala Ala Val Gln Cys Ile Ser Ala
345                 350                 355                 360

gaa gtt cag aga aat tcg agt acc cag gga act gcg gag gca ctc gca    13570
Glu Val Gln Arg Asn Ser Ser Thr Gln Gly Thr Ala Glu Ala Leu Ala
                    365                 370                 375

gaa tca gtc gtg aag ccc acg aga cga agg tca tct cag tgg aag aag    13618
Glu Ser Val Val Lys Pro Thr Arg Arg Arg Ser Ser Gln Trp Lys Lys
                380                 385                 390

tcg aca cac ccc cta tca gaa gta gca gta cac aac aag cca agc gat    13666
Ser Thr His Pro Leu Ser Glu Val Ala Val His Asn Lys Pro Ser Asp
        395                 400                 405

tgc tgg att gtt gta aaa aac aag gtg tat gat gtt tcc aat ttt gcg    13714
Cys Trp Ile Val Val Lys Asn Lys Val Tyr Asp Val Ser Asn Phe Ala
        410                 415                 420

gac gag cat ccc gga gga tca gtt att agt act tat ttt gga cga gac    13762
Asp Glu His Pro Gly Gly Ser Val Ile Ser Thr Tyr Phe Gly Arg Asp
425                 430                 435                 440

ggc aca gat gtt ttc tct agt ttt cat gca gct tct aca tgg aaa att    13810
Gly Thr Asp Val Phe Ser Ser Phe His Ala Ala Ser Thr Trp Lys Ile
                    445                 450                 455

ctt caa gac ttt tac att ggt gac gtg gag agg gtg gag ccg act cca    13858
Leu Gln Asp Phe Tyr Ile Gly Asp Val Glu Arg Val Glu Pro Thr Pro
                    460                 465                 470

gag ctg ctg aaa gat ttc cga gaa atg aga gct ctt ttc ctg agg gag    13906
Glu Leu Leu Lys Asp Phe Arg Glu Met Arg Ala Leu Phe Leu Arg Glu
            475                 480                 485

caa ctt ttc aaa agt tcg aaa ttg tac tat gtt atg aag ctg ctc acg    13954
Gln Leu Phe Lys Ser Ser Lys Leu Tyr Tyr Val Met Lys Leu Leu Thr
        490                 495                 500

aat gtt gct att ttt gct gcg agc att gca ata ata tgt tgg agc aag    14002
Asn Val Ala Ile Phe Ala Ala Ser Ile Ala Ile Ile Cys Trp Ser Lys
505                 510                 515                 520

act att tca gcg gtt ttg gct tca gct tgt atg atg gct ctg tgt ttc    14050
Thr Ile Ser Ala Val Leu Ala Ser Ala Cys Met Met Ala Leu Cys Phe
                    525                 530                 535

caa cag tgc gga tgg cta tcc cat gat ttt ctc cac aat cag gtg ttt    14098
Gln Gln Cys Gly Trp Leu Ser His Asp Phe Leu His Asn Gln Val Phe
                    540                 545                 550

gag aca cgc tgg ctt aat gaa gtt gtc ggg tat gtg atc ggc aac gcc    14146
Glu Thr Arg Trp Leu Asn Glu Val Val Gly Tyr Val Ile Gly Asn Ala
                555                 560                 565
```

```
gtt ctg ggg ttt agt aca ggg tgg tgg aag gag aag cat aac ctt cat    14194
Val Leu Gly Phe Ser Thr Gly Trp Trp Lys Glu Lys His Asn Leu His
    570             575             580

cat gct gct cca aat gaa tgc gat cag act tac caa cca att gat gaa    14242
His Ala Ala Pro Asn Glu Cys Asp Gln Thr Tyr Gln Pro Ile Asp Glu
585             590             595             600

gat att gat act ctc ccc ctc att gcc tgg agc aag gac ata ctg gcc    14290
Asp Ile Asp Thr Leu Pro Leu Ile Ala Trp Ser Lys Asp Ile Leu Ala
            605             610             615

aca gtt gag aat aag aca ttc ttg cga atc ctc caa tac cag cat ctg    14338
Thr Val Glu Asn Lys Thr Phe Leu Arg Ile Leu Gln Tyr Gln His Leu
            620             625             630

ttc ttc atg ggt ctg tta ttt ttc gcc cgt ggt agt tgg ctc ttt tgg    14386
Phe Phe Met Gly Leu Leu Phe Phe Ala Arg Gly Ser Trp Leu Phe Trp
        635             640             645

agc tgg aga tat acc tct aca gca gtg ctc tca cct gtc gac agg ttg    14434
Ser Trp Arg Tyr Thr Ser Thr Ala Val Leu Ser Pro Val Asp Arg Leu
    650             655             660

ttg gag aag gga act gtt ctg ttt cac tac ttt tgg ttc gtc ggg aca    14482
Leu Glu Lys Gly Thr Val Leu Phe His Tyr Phe Trp Phe Val Gly Thr
665             670             675             680

gcg tgc tat ctt ctc cct ggt tgg aag cca tta gta tgg atg gcg gtg    14530
Ala Cys Tyr Leu Leu Pro Gly Trp Lys Pro Leu Val Trp Met Ala Val
            685             690             695

act gag ctc atg tcc ggc atg ctg ctg ggc ttt gta ttt gta ctt agc    14578
Thr Glu Leu Met Ser Gly Met Leu Leu Gly Phe Val Phe Val Leu Ser
            700             705             710

cac aat ggg atg gag gtt tat aat tcg tct aaa gaa ttc gtg agt gca    14626
His Asn Gly Met Glu Val Tyr Asn Ser Ser Lys Glu Phe Val Ser Ala
        715             720             725

cag atc gta tcc aca cgg gat atc aaa gga aac ata ttc aac gac tgg    14674
Gln Ile Val Ser Thr Arg Asp Ile Lys Gly Asn Ile Phe Asn Asp Trp
    730             735             740

ttc act ggt ggc ctt aac agg caa ata gag cat cat ctt ttc cca aca    14722
Phe Thr Gly Gly Leu Asn Arg Gln Ile Glu His His Leu Phe Pro Thr
745             750             755             760

atg ccc agg cat aat tta aac aaa ata gca cct aga gtg gag gtg ttc    14770
Met Pro Arg His Asn Leu Asn Lys Ile Ala Pro Arg Val Glu Val Phe
            765             770             775

tgt aag aaa cac ggt ctg gtg tac gaa gac gta tct att gct acc ggc    14818
Cys Lys Lys His Gly Leu Val Tyr Glu Asp Val Ser Ile Ala Thr Gly
            780             785             790
```

```
act tgc aag gtt ttg aaa gca ttg aag gaa gtc gcg gag gct gcg gca    14866
Thr Cys Lys Val Leu Lys Ala Leu Lys Glu Val Ala Glu Ala Ala Ala
        795             800             805

gag cag cat gct acc acc agt taa gctagcgtta accctgcttt aatgagatat   14920
Glu Gln His Ala Thr Thr Ser
    810             815

gcgagacgcc tatgatcgca tgatatttgc tttcaattct gttgtgcacg ttgtaaaaaa   14980

cctgagcatg tgtagctcag atccttaccg ccggtttcgg ttcattctaa tgaatatatc   15040

acccgttact atcgtatttt tatgaataat attctccgtt caatttactg attgtccgtc   15100

gagcaaattt acacattgcc actaaacgtc taaacccttg taatttgttt ttgttttact   15160

atgtgtgtta tgtatttgat ttgcgataaa tttttatatt tggtactaaa tttataacac   15220

cttttatgct aacgtttgcc aacacttagc aatttgcaag ttgattaatt gattctaaat   15280

tattttttgtc ttctaaatac atatactaat caactggaaa tgtaaatatt tgctaatatt   15340

tctactatag gagaattaaa gtgagtgaat atggtaccac aaggtttgga gatttaattg   15400

ttgcaatgct gcatggatgg catatacacc aaacattcaa taattcttga ggataataat   15460

ggtaccacac aagatttgag gtgcatgaac gtcacgtgga caaaaggttt agtaattttt   15520

caagacaaca atgttaccac acacaagttt tgaggtgcat gcatggatgc cctgtggaaa   15580

gtttaaaaat attttggaaa tgatttgcat ggaagccatg tgtaaaacca tgacatccac   15640

ttggaggatg caataatgaa gaaaactaca aatttacatg caactagtta tgcatgtagt   15700

ctatataatg aggattttgc aatactttca ttcatacaca ctcactaagt tttacacgat   15760

tataatttct tcatagccag cagatctaaa atg gct ccg gat gcg gat aag ctt   15814
                                     Met Ala Pro Asp Ala Asp Lys Leu
                                                     820

cga caa cgc cag acg act gcg gta gcg aag cac aat gct gct acc ata   15862
Arg Gln Arg Gln Thr Thr Ala Val Ala Lys His Asn Ala Ala Thr Ile
    825             830             835

tcg acg cag gaa cgc ctt tgc agt ctg tct tcg ctc aaa ggc gaa gaa   15910
Ser Thr Gln Glu Arg Leu Cys Ser Leu Ser Ser Leu Lys Gly Glu Glu
840             845             850             855

gtc tgc atc gac gga atc atc tat gac ctc caa tca ttc gat cat ccc   15958
Val Cys Ile Asp Gly Ile Ile Tyr Asp Leu Gln Ser Phe Asp His Pro
                860             865             870

ggg ggt gaa acg atc aaa atg ttt ggt ggc aac gat gtc act gta cag   16006
Gly Gly Glu Thr Ile Lys Met Phe Gly Gly Asn Asp Val Thr Val Gln
            875             880             885

tac aag atg att cac ccg tac cat acc gag aag cat ttg gaa aag atg   16054
```

Tyr Lys Met Ile His Pro Tyr His Thr Glu Lys His Leu Glu Lys Met
    890   895   900

aag cgt gtc ggc aag gtg acg gat ttc gtc tgc gag tac aag ttc gat   16102
Lys Arg Val Gly Lys Val Thr Asp Phe Val Cys Glu Tyr Lys Phe Asp
    905   910   915

acc gaa ttt gaa cgc gaa atc aaa cga gaa gtc ttc aag att gtg cga   16150
Thr Glu Phe Glu Arg Glu Ile Lys Arg Glu Val Phe Lys Ile Val Arg
920   925   930   935

cga ggc aag gat ttc ggt act ttg gga tgg ttc ttc cgt gcg ttt tgc   16198
Arg Gly Lys Asp Phe Gly Thr Leu Gly Trp Phe Phe Arg Ala Phe Cys
         940   945   950

tac att gcc att ttc ttc tac ctg cag tac cat tgg gtc acc acg gga   16246
Tyr Ile Ala Ile Phe Phe Tyr Leu Gln Tyr His Trp Val Thr Thr Gly
         955   960   965

acc tct tgg ctg ctg gcc gtg gcc tac gga atc tcc caa gcg atg att   16294
Thr Ser Trp Leu Leu Ala Val Ala Tyr Gly Ile Ser Gln Ala Met Ile
    970   975   980

ggc atg aat gtc cag cac gat gcc aac cac ggg gcc acc tcc aag cgt   16342
Gly Met Asn Val Gln His Asp Ala Asn His Gly Ala Thr Ser Lys Arg
    985   990   995

ccc tgg gtc aac gac atg cta ggc ctc ggt gcg gat ttt att ggt   16387
Pro Trp Val Asn Asp Met Leu Gly Leu Gly Ala Asp Phe Ile Gly
1000   1005   1010

ggt tcc aag tgg ctc tgg cag gaa caa cac tgg acc cac cac gct   16432
Gly Ser Lys Trp Leu Trp Gln Glu Gln His Trp Thr His His Ala
1015   1020   1025

tac acc aat cac gcc gag atg gat ccc gat agc ttt ggt gcc gaa   16477
Tyr Thr Asn His Ala Glu Met Asp Pro Asp Ser Phe Gly Ala Glu
1030   1035   1040

cca atg ctc cta ttc aac gac tat ccc ttg gat cat ccc gct cgt   16522
Pro Met Leu Leu Phe Asn Asp Tyr Pro Leu Asp His Pro Ala Arg
1045   1050   1055

acc tgg cta cat cgc ttt caa gca ttc ttt tac atg ccc gtc ttg   16567
Thr Trp Leu His Arg Phe Gln Ala Phe Phe Tyr Met Pro Val Leu
1060   1065   1070

gct gga tac tgg ttg tcc gct gtc ttc aat cca caa att ctt gac   16612
Ala Gly Tyr Trp Leu Ser Ala Val Phe Asn Pro Gln Ile Leu Asp
1075   1080   1085

ctc cag caa cgc ggc gca ctt tcc gtc ggt atc cgt ctc gac aac   16657
Leu Gln Gln Arg Gly Ala Leu Ser Val Gly Ile Arg Leu Asp Asn
1090   1095   1100

gct ttc att cac tcg cga cgc aag tat gcg gtt ttc tgg cgg gct   16702
Ala Phe Ile His Ser Arg Arg Lys Tyr Ala Val Phe Trp Arg Ala

1105   1110   1115

```
gtg  tac  att  gcg  gtg  aac  gtg  att  gct  ccg  ttt  tac  aca  aac  tcc      16747
Val  Tyr  Ile  Ala  Val  Asn  Val  Ile  Ala  Pro  Phe  Tyr  Thr  Asn  Ser
1120                     1125                     1130

ggc  ctc  gaa  tgg  tcc  tgg  cgt  gtc  ttt  gga  aac  atc  atg  ctc  atg      16792
Gly  Leu  Glu  Trp  Ser  Trp  Arg  Val  Phe  Gly  Asn  Ile  Met  Leu  Met
1135                     1140                     1145

ggt  gtg  gcg  gaa  tcg  ctc  gcg  ctg  gcg  gtc  ctg  ttt  tcg  ttg  tcg      16837
Gly  Val  Ala  Glu  Ser  Leu  Ala  Leu  Ala  Val  Leu  Phe  Ser  Leu  Ser
1150                     1155                     1160

cac  aat  ttc  gaa  tcc  gcg  gat  cgc  gat  ccg  acc  gcc  cca  ctg  aaa      16882
His  Asn  Phe  Glu  Ser  Ala  Asp  Arg  Asp  Pro  Thr  Ala  Pro  Leu  Lys
1165                     1170                     1175

aag  acg  gga  gaa  cca  gtc  gac  tgg  ttc  aag  aca  cag  gtc  gaa  act      16927
Lys  Thr  Gly  Glu  Pro  Val  Asp  Trp  Phe  Lys  Thr  Gln  Val  Glu  Thr
1180                     1185                     1190

tcc  tgc  act  tac  ggt  gga  ttc  ctt  tcc  ggt  tgc  ttc  acg  gga  ggt      16972
Ser  Cys  Thr  Tyr  Gly  Gly  Phe  Leu  Ser  Gly  Cys  Phe  Thr  Gly  Gly
1195                     1200                     1205

ctc  aac  ttt  cag  gtt  gaa  cac  cac  ttg  ttc  cca  cgc  atg  agc  agc      17017
Leu  Asn  Phe  Gln  Val  Glu  His  His  Leu  Phe  Pro  Arg  Met  Ser  Ser
1210                     1215                     1220

gct  tgg  tat  ccc  tac  att  gcc  ccc  aag  gtc  cgc  gaa  att  tgc  gcc      17062
Ala  Trp  Tyr  Pro  Tyr  Ile  Ala  Pro  Lys  Val  Arg  Glu  Ile  Cys  Ala
1225                     1230                     1235

aaa  cac  ggc  gtc  cac  tac  gcc  tac  tac  ccg  tgg  atc  cac  caa  aac      17107
Lys  His  Gly  Val  His  Tyr  Ala  Tyr  Tyr  Pro  Trp  Ile  His  Gln  Asn
1240                     1245                     1250

ttt  ctc  tcc  acc  gtc  cgc  tac  atg  cac  gcg  gcc  ggg  acc  ggt  gcc      17152
Phe  Leu  Ser  Thr  Val  Arg  Tyr  Met  His  Ala  Ala  Gly  Thr  Gly  Ala
1255                     1260                     1265

aac  tgg  cgc  cag  atg  gcc  aga  gaa  aat  ccc  ttg  acc  gga  cgg  gcg      17197
Asn  Trp  Arg  Gln  Met  Ala  Arg  Glu  Asn  Pro  Leu  Thr  Gly  Arg  Ala
1270                     1275                     1280

taa  agatctgccg  gcatcgatcc  cgggccatgg  cctgctttaa  tgagatatgc      17250

gagacgccta  tgatcgcatg  atatttgctt  tcaattctgt  tgtgcacgtt  gtaaaaaacc      17310

tgagcatgtg  tagctcagat  ccttaccgcc  ggtttcggtt  cattctaatg  aatatatcac      17370

ccgttactat  cgtattttta  tgaataatat  tctccgttca  atttactgat  tgtccgtcga      17430

cgagctcggc  gcgcctctag  aggatcgatg  aattcagatc  ggctgagtgg  ctccttcaac      17490

gttgcggttc  tgtcagttcc  aaacgtaaaa  cggcttgtcc  cgcgtcatcg  gcggggggtca      17550
```

taacgtgact cccttaattc tccgctcatg atcagattgt cgtttcccgc cttcagttta 17610

aactatcagt gtttgacagg atatattggc gggtaaacct aagagaaaag agcgtttatt 17670

agaataatcg gatatttaaa agggcgtgaa aaggtttatc cttcgtccat ttgtatgtgc 17730

atgccaacca cagggttccc ca 17752

<210> 20
<211> 290
<212> PRT
<213> Phaeodactylum tricornutum, Physcomitrella patens

<400> 20

Met Glu Val Val Glu Arg Phe Tyr Gly Glu Leu Asp Gly Lys Val Ser
1   5   10   15

Gln Gly Val Asn Ala Leu Leu Gly Ser Phe Gly Val Glu Leu Thr Asp
            20   25   30

Thr Pro Thr Thr Lys Gly Leu Pro Leu Val Asp Ser Pro Thr Pro Ile
        35   40   45

Val Leu Gly Val Ser Val Tyr Leu Thr Ile Val Ile Gly Gly Leu Leu
    50   55   60

Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg Ala Ser Glu Pro Phe Leu
65   70   75   80

Leu Gln Ala Leu Val Leu Val His Asn Leu Phe Cys Phe Ala Leu Ser
                85   90   95

Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln Ala Ile Thr Trp Arg Tyr
            100   105   110

Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys His Lys Glu Met Ala Ile
            115   120   125

Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr Val Glu Phe Met Asp Thr
        130   135   140

Val Ile Met Ile Leu Lys Arg Ser Thr Arg Gln Ile Ser Phe Leu His
145   150   155   160

Val Tyr His His Ser Ser Ile Ser Leu Ile Trp Trp Ala Ile Ala His
           165   170   175

His Ala Pro Gly Gly Glu Ala Tyr Trp Ser Ala Ala Leu Asn Ser Gly
          180   185   190

Val His Val Leu Met Tyr Ala Tyr Tyr Phe Leu Ala Ala Cys Leu Arg
          195   200   205

Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu Phe Trp Gly Arg Tyr Leu
     210   215   220

Thr Gln Phe Gln Met Phe Gln Phe Met Leu Asn Leu Val Gln Ala Tyr
225   230   235   240

Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro Gln Trp Leu Ile Lys Ile
          245   250   255

Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe Leu Phe Gly Asn Phe Tyr
          260   265   270

Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly Lys Gln Lys Gly Ala Lys
          275   280   285

Thr Glu
   290

<210> 21
<211> 525
<212> PRT
<213> Phaeodactylum tricornutum, Physcomitrella patens

<400> 21

Met Val Phe Ala Gly Gly Gly Leu Gln Gln Gly Ser Leu Glu Glu Asn
1   5   10   15

Ile Asp Val Glu His Ile Ala Ser Met Ser Leu Phe Ser Asp Phe Phe
          20   25   30

Ser Tyr Val Ser Ser Thr Val Gly Ser Trp Ser Val His Ser Ile Gln
          35   40   45

Pro Leu Lys Arg Leu Thr Ser Lys Lys Arg Val Ser Glu Ser Ala Ala
     50   55   60

Val Gln Cys Ile Ser Ala Glu Val Gln Arg Asn Ser Ser Thr Gln Gly
65   70   75   80

Thr Ala Glu Ala Leu Ala Glu Ser Val Val Lys Pro Thr Arg Arg Arg
                      85   90   95

Ser Ser Gln Trp Lys Lys Ser Thr His Pro Leu Ser Glu Val Ala Val
              100   105   110

His Asn Lys Pro Ser Asp Cys Trp Ile Val Val Lys Asn Lys Val Tyr
          115   120   125

Asp Val Ser Asn Phe Ala Asp Glu His Pro Gly Gly Ser Val Ile Ser
    130   135   140

Thr Tyr Phe Gly Arg Asp Gly Thr Asp Val Phe Ser Ser Phe His Ala
145   150   155   160

Ala Ser Thr Trp Lys Ile Leu Gln Asp Phe Tyr Ile Gly Asp Val Glu
              165   170   175

Arg Val Glu Pro Thr Pro Glu Leu Leu Lys Asp Phe Arg Glu Met Arg
              180   185   190

Ala Leu Phe Leu Arg Glu Gln Leu Phe Lys Ser Ser Lys Leu Tyr Tyr
          195   200   205

Val Met Lys Leu Leu Thr Asn Val Ala Ile Phe Ala Ala Ser Ile Ala
    210   215   220

Ile Ile Cys Trp Ser Lys Thr Ile Ser Ala Val Leu Ala Ser Ala Cys
225   230   235   240

Met Met Ala Leu Cys Phe Gln Gln Cys Gly Trp Leu Ser His Asp Phe
              245   250   255

Leu His Asn Gln Val Phe Glu Thr Arg Trp Leu Asn Glu Val Val Gly
          260   265   270

Tyr Val Ile Gly Asn Ala Val Leu Gly Phe Ser Thr Gly Trp Trp Lys
    275   280   285

Glu Lys His Asn Leu His His Ala Ala Pro Asn Glu Cys Asp Gln Thr
    290   295   300

Tyr Gln Pro Ile Asp Glu Asp Ile Asp Thr Leu Pro Leu Ile Ala Trp
305   310   315   320

Ser Lys Asp Ile Leu Ala Thr Val Glu Asn Lys Thr Phe Leu Arg Ile
            325   330   335

Leu Gln Tyr Gln His Leu Phe Phe Met Gly Leu Leu Phe Phe Ala Arg
            340   345   350

Gly Ser Trp Leu Phe Trp Ser Trp Arg Tyr Thr Ser Thr Ala Val Leu
            355   360   365

Ser Pro Val Asp Arg Leu Leu Glu Lys Gly Thr Val Leu Phe His Tyr
    370   375   380

Phe Trp Phe Val Gly Thr Ala Cys Tyr Leu Leu Pro Gly Trp Lys Pro
385   390   395   400

Leu Val Trp Met Ala Val Thr Glu Leu Met Ser Gly Met Leu Leu Gly
            405   410   415

Phe Val Phe Val Leu Ser His Asn Gly Met Glu Val Tyr Asn Ser Ser
            420   425   430

Lys Glu Phe Val Ser Ala Gln Ile Val Ser Thr Arg Asp Ile Lys Gly
    435   440   445

Asn Ile Phe Asn Asp Trp Phe Thr Gly Gly Leu Asn Arg Gln Ile Glu
    450   455   460

His His Leu Phe Pro Thr Met Pro Arg His Asn Leu Asn Lys Ile Ala
465   470   475   480

Pro Arg Val Glu Val Phe Cys Lys Lys His Gly Leu Val Tyr Glu Asp
            485   490   495

Val Ser Ile Ala Thr Gly Thr Cys Lys Val Leu Lys Ala Leu Lys Glu
            500   505   510

Val Ala Glu Ala Ala Ala Glu Gln His Ala Thr Thr Ser
     515   520   525

<210> 22
<211> 469
<212> PRT
<213> Phaeodactylum tricornutum, Physcomitrella patens

<400> 22

Met Ala Pro Asp Ala Asp Lys Leu Arg Gln Arg Gln Thr Thr Ala Val
1   5   10   15

Ala Lys His Asn Ala Ala Thr Ile Ser Thr Gln Glu Arg Leu Cys Ser
       20   25   30

Leu Ser Ser Leu Lys Gly Glu Glu Val Cys Ile Asp Gly Ile Ile Tyr
     35   40   45

Asp Leu Gln Ser Phe Asp His Pro Gly Gly Glu Thr Ile Lys Met Phe
   50   55   60

Gly Gly Asn Asp Val Thr Val Gln Tyr Lys Met Ile His Pro Tyr His
65   70   75   80

Thr Glu Lys His Leu Glu Lys Met Lys Arg Val Gly Lys Val Thr Asp
       85   90   95

Phe Val Cys Glu Tyr Lys Phe Asp Thr Glu Phe Glu Arg Glu Ile Lys
      100   105   110

Arg Glu Val Phe Lys Ile Val Arg Arg Gly Lys Asp Phe Gly Thr Leu
     115   120   125

Gly Trp Phe Phe Arg Ala Phe Cys Tyr Ile Ala Ile Phe Phe Tyr Leu
   130   135   140

Gln Tyr His Trp Val Thr Thr Gly Thr Ser Trp Leu Leu Ala Val Ala
145   150   155   160

Tyr Gly Ile Ser Gln Ala Met Ile Gly Met Asn Val Gln His Asp Ala
       165   170   175

Asn His Gly Ala Thr Ser Lys Arg Pro Trp Val Asn Asp Met Leu Gly
       180   185   190

Leu Gly Ala Asp Phe Ile Gly Gly Ser Lys Trp Leu Trp Gln Glu Gln
    195   200   205

His Trp Thr His His Ala Tyr Thr Asn His Ala Glu Met Asp Pro Asp
    210   215   220

Ser Phe Gly Ala Glu Pro Met Leu Leu Phe Asn Asp Tyr Pro Leu Asp
225   230   235   240

His Pro Ala Arg Thr Trp Leu His Arg Phe Gln Ala Phe Phe Tyr Met
         245   250   255

Pro Val Leu Ala Gly Tyr Trp Leu Ser Ala Val Phe Asn Pro Gln Ile
         260   265   270

Leu Asp Leu Gln Gln Arg Gly Ala Leu Ser Val Gly Ile Arg Leu Asp
         275   280   285

Asn Ala Phe Ile His Ser Arg Arg Lys Tyr Ala Val Phe Trp Arg Ala
    290   295   300

Val Tyr Ile Ala Val Asn Val Ile Ala Pro Phe Tyr Thr Asn Ser Gly
305   310   315   320

Leu Glu Trp Ser Trp Arg Val Phe Gly Asn Ile Met Leu Met Gly Val
         325   330   335

Ala Glu Ser Leu Ala Leu Ala Val Leu Phe Ser Leu Ser His Asn Phe
         340   345   350

Glu Ser Ala Asp Arg Asp Pro Thr Ala Pro Leu Lys Lys Thr Gly Glu

        355   360   365

Pro Val Asp Trp Phe Lys Thr Gln Val Glu Thr Ser Cys Thr Tyr Gly
    370   375   380

Gly Phe Leu Ser Gly Cys Phe Thr Gly Gly Leu Asn Phe Gln Val Glu
385   390   395   400

His His Leu Phe Pro Arg Met Ser Ser Ala Trp Tyr Pro Tyr Ile Ala
         405   410   415

Pro Lys Val Arg Glu Ile Cys Ala Lys His Gly Val His Tyr Ala Tyr
        420   425   430

Tyr Pro Trp Ile His Gln Asn Phe Leu Ser Thr Val Arg Tyr Met His
        435   440   445

Ala Ala Gly Thr Gly Ala Asn Trp Arg Gln Met Ala Arg Glu Asn Pro
        450   455   460

Leu Thr Gly Arg Ala
465

<210>   23
<211>   26
<212>   DNA
<213>   artificial sequence

<220>
<221>   misc_feature
<222>   (1)..(26)
<223>   polylinker sequence

<400>   23
gaattcggcg cgccgagctc ctcgag   26

<210>   24
<211>   265
<212>   DNA
<213>   artificial sequence

<220>
<221>   misc_feature
<222>   (1)..(265)
<223>   polylinker-terminator-polylinkers sequence

<400>   24
ccaccgcggt gggcggccgc ctgcagtcta gaaggcctcc tgctttaatg agatatgcga   60

gacgcctatg atcgcatgat atttgctttc aattctgttg tgcacgttgt aaaaaacctg   120

agcatgtgta gctcagatcc ttaccgccgg tttcggttca ttctaatgaa tatatcaccc   180

gttactatcg tattttatg aataatattc tccgttcaat ttactgattg tccgtcgacg   240

aattcgagct cggcgcgcca agctt   265

<210>   25

```
<211>   257
<212>   DNA
<213>   artificial sequence

<220>
<221>   misc_feature
<222>   (1)..(257)
<223>   polylinker-terminator-polylinkers sequence



<400>  25
ggatccgata tcgggcccgc tagcgttaac cctgctttaa tgagatatgc gagacgccta      60

tgatcgcatg atatttgctt tcaattctgt tgtgcacgtt gtaaaaaacc tgagcatgtg     120

tagctcagat ccttaccgcc ggtttcggtt cattctaatg aatatatcac ccgttactat     180

cgtattttta tgaataatat tctccgttca attactgat  tgtccgtcga cgaattcgag     240

ctcggcgcgc caagctt                                                     257


<210>   26
<211>   5410
<212>   DNA
<213>   artificial sequence

<220>
<221>   misc_feature
<222>   (1)..(5410)
<223>   plant expression vector with one promoter-terminator expression cassette

<400>  26
ttttggaaat gatttgcatg gaagccatgt gtaaaaccat gacatccact tggaggatgc      60

aataatgaag aaaactacaa atttacatgc aactagttat gcatgtagtc tatataatga     120

ggattttgca atactttcat tcatacacac tcactaagtt ttacacgatt ataatttctt     180

catagccagc ggatccgata tcgggcccgc tagcgttaac cctgctttaa tgagatatgc     240

gagacgccta tgatcgcatg atatttgctt tcaattctgt tgtgcacgtt gtaaaaaacc     300

tgagcatgtg tagctcagat ccttaccgcc ggtttcggtt cattctaatg aatatatcac     360

ccgttactat cgtattttta tgaataatat tctccgttca attactgat  tgtccgtcga     420

gcaaatttac acattgccac taaacgtcta aacccttgta atttgttttt gttttactat     480

gtgtgttatg tatttgattt gcgataaatt tttatatttg gtactaaatt tataacacct     540

tttatgctaa cgtttgccaa cacttagcaa tttgcaagtt gattaattga ttctaaatta     600

tttttgtctt ctaaatacat atactaatca actggaaatg taaatatttg ctaatatttc     660

tactatagga gaattaaagt gagtgaatat ggtaccacaa ggtttggaga tttaattgtt     720
```

```
gcaatgctgc atggatggca tatacaccaa acattcaata attcttgagg ataataatgg     780

taccacacaa gatttgaggt gcatgaacgt cacgtggaca aaaggtttag taattttca      840

agacaacaat gttaccacac acaagttttg aggtgcatgc atggatgccc tgtggaaagt     900

ttaaaaatat tttggaaatg atttgcatgg aagccatgtg taaaaccatg acatccactt     960

ggaggatgca ataatgaaga aaactacaaa tttacatgca actagttatg catgtagtct    1020

atataatgag gattttgcaa tactttcatt catacacact cactaagttt tacacgatta    1080

taatttcttc atagccagca gatctgccgg catcgatccc gggccatggc ctgctttaat    1140

gagatatgcg agacgcctat gatcgcatga tatttgcttt caattctgtt gtgcacgttg    1200

taaaaaacct gagcatgtgt agctcagatc cttaccgccg gtttcggttc attctaatga    1260

atatatcacc cgttactatc gtattttat gaataatatt ctccgttcaa tttactgatt     1320

gtccgtcgac gagctcggcg cgccaagctt ggcgtaatca tggtcatagc tgtttcctgt    1380

gtgaaattgt tatccgctca caattccaca caacatacga gccggaagca taaagtgtaa    1440

agcctggggt gcctaatgag tgagctaact cacattaatt gcgttgcgct cactgcccgc    1500

tttccagtcg ggaaacctgt cgtgccagct gcattaatga atcggccaac gcgcggggag    1560

aggcggtttg cgtattgggc gctcttccgc ttcctcgctc actgactcgc tgcgctcggt    1620

cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt tatccacaga    1680

atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg    1740

taaaaaggcc gcgttgctgg cgttttttcca taggctccgc cccctgacg agcatcacaa    1800

aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt    1860

tcccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct     1920

gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct    1980

cagttcggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc    2040

cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt    2100

atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc    2160

tacagagttc ttgaagtggt ggcctaacta cggctacact agaaggacag tatttggtat    2220

ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa    2280

acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa    2340
```

```
aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga   2400

aaactcacgt taagggattt tggtcatgag attatcaaaa aggatcttca cctagatcct   2460

tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa cttggtctga   2520

cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcatc   2580

catagttgcc tgactccccg tcgtgtagat aactacgata cgggagggct taccatctgg   2640

ccccagtgct gcaatgatac cgcgagaccc acgctcaccg gctccagatt tatcagcaat   2700

aaaccagcca gccggaaggg ccgagcgcag aagtggtcct gcaactttat ccgcctccat   2760

ccagtctatt aattgttgcc gggaagctag agtaagtagt tcgccagtta atagtttgcg   2820

caacgttgtt gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg gtatggcttc   2880

attcagctcc ggttcccaac gatcaaggcg agttacatga tcccccatgt tgtgcaaaaa   2940

agcggttagc tccttcggtc ctccgatcgt tgtcagaagt aagttggccg cagtgttatc   3000

actcatggtt atggcagcac tgcataattc tcttactgtc atgccatccg taagatgctt   3060

ttctgtgact ggtgagtact caaccaagtc attctgagaa tagtgtatgc ggcgaccgag   3120

ttgctcttgc ccggcgtcaa tacgggataa taccgcgcca catagcagaa ctttaaaagt   3180

gctcatcatt ggaaaacgtt cttcggggcg aaaactctca aggatcttac cgctgttgag   3240

atccagttcg atgtaaccca ctcgtgcacc caactgatct tcagcatctt ttactttcac   3300

cagcgtttct gggtgagcaa aaacaggaag gcaaaatgcc gcaaaaaagg gaataagggc   3360

gacacggaaa tgttgaatac tcatactctt cctttttcaa tattattgaa gcatttatca   3420

gggttattgt ctcatgagcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg   3480

ggttccgcgc acatttcccc gaaaagtgcc acctgacgtc taagaaacca ttattatcat   3540

gacattaacc tataaaaata ggcgtatcac gaggcccttt cgtctcgcgc gtttcggtga   3600

tgacggtgaa aacctctgac acatgcagct cccggagacg gtcacagctt gtctgtaagc   3660

ggatgccggg agcagacaag cccgtcaggg cgcgtcagcg ggtgttggcg ggtgtcgggg   3720

ctggcttaac tatgcggcat cagagcagat tgtactgaga gtgcaccata tgcggtgtga   3780

aataccgcac agatgcgtaa ggagaaaata ccgcatcagg cgccattcgc cattcaggct   3840

gcgcaactgt tgggaagggc gatcggtgcg ggcctcttcg ctattacgcc agctggcgaa   3900

agggggatgt gctgcaaggc gattaagttg ggtaacgcca gggttttccc agtcacgacg   3960

ttgtaaaacg acggccagtg aattcggcgc gccgagctcc tcgagcaaat ttacacattg   4020

ccactaaacg tctaaaccct tgtaatttgt ttttgtttta ctatgtgtgt tatgtatttg   4080
```

```
atttgcgata aattttata tttggtacta aatttataac acctttatg ctaacgtttg      4140

ccaacactta gcaatttgca agttgattaa ttgattctaa attatttttg tcttctaaat      4200

acatatacta atcaactgga aatgtaaata tttgctaata tttctactat aggagaatta      4260

aagtgagtga atatggtacc acaaggtttg gagatttaat tgttgcaatg ctgcatggat      4320

ggcatataca ccaaacattc aataattctt gaggataata atggtaccac acaagatttg      4380

aggtgcatga acgtcacgtg gacaaaaggt ttagtaattt ttcaagacaa caatgttacc      4440

acacacaagt tttgaggtgc atgcatggat gccctgtgga aagtttaaaa atattttgga      4500

aatgatttgc atggaagcca tgtgtaaaac catgacatcc acttggagga tgcaataatg      4560

aagaaaacta caaatttaca tgcaactagt tatgcatgta gtctatataa tgaggatttt      4620


gcaatacttt cattcataca cactcactaa gttttacacg attataattt cttcatagcc      4680

agcccaccgc ggtgggcggc cgcctgcagt ctagaaggcc tcctgcttta atgagatatg      4740

cgagacgcct atgatcgcat gatatttgct ttcaattctg ttgtgcacgt tgtaaaaaac      4800

ctgagcatgt gtagctcaga tccttaccgc cggtttcggt tcattctaat gaatatatca      4860

cccgttacta tcgtattttt atgaataata ttctccgttc aatttactga ttgtccgtcg      4920

agcaaattta cacattgcca ctaaacgtct aaacccttgt aatttgtttt tgttttacta      4980

tgtgtgttat gtatttgatt tgcgataaat ttttatattt ggtactaaat ttataacacc      5040

ttttatgcta acgtttgcca acacttagca atttgcaagt tgattaattg attctaaatt      5100

attttttgtct tctaaataca tatactaatc aactggaaat gtaaatattt gctaatattt      5160

ctactatagg agaattaaag tgagtgaata tggtaccaca aggtttggag atttaattgt      5220

tgcaatgctg catggatggc atatacacca aacattcaat aattcttgag gataataatg      5280

gtaccacaca agatttgagg tgcatgaacg tcacgtggac aaaaggttta gtaatttttc      5340

aagacaacaa tgttaccaca cacaagtttt gaggtgcatg catggatgcc ctgtggaaag      5400

tttaaaaata                                                              5410
```

```
<210>  27
<211>  12093
<212>  DNA
<213>  artificial sequence

<220>
<221>  misc_feature
<222>  (1)..(12093)
```

<223> plant expression vector with one promoter-terminator expression cassette

<400> 27

```
gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg atccgacagc      60

gcgcccagca caggtgcgca ggcaaattgc accaacgcat acagcgccag cagaatgcca     120

tagtgggcgg tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccgg cagcaccggc     180

ataatcaggc cgatgccgac agcgtcgagc gcgacagtgc tcagaattac gatcaggggt     240

atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg aacgcgcgga     300

ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa gtgtcaagca     360

tgacaaagtt gcagccgaat acagtgatcc gtgccgccct ggacctgttg aacgaggtcg     420

gcgtagacgg tctgacgaca cgcaaactgg cggaacggtt ggggggttcag cagccggcgc     480

tttactggca cttcaggaac aagcgggcgc tgctcgacgc actggccgaa gccatgctgg     540

cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca tttctgatcg     600

ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg cgcatccatg     660

ccggcacgcg accgggcgca ccgcagatgg aaacggccga cgcgcagctt cgcttcctct     720

gcgaggcggg tttttcggcc ggggacgccg tcaatgcgct gatgacaatc agctacttca     780

ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag cgcggcggca     840

ccgttgaaca ggctccgctc tcgccgctgt tgcgggccgc gatagacgcc ttcgacgaag     900

ccggtccgga cgcagcgttc gagcagggac tcgcggtgat tgtcgatgga ttggcgaaaa     960

ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc    1020

tgccggagcg caacccactc actacagcag agccatgtag acaacatccc ctcccccttt    1080

ccaccgcgtc agacgcccgt agcagcccgc tacgggcttt ttcatgccct gccctagcgt    1140

ccaagcctca cggccgcgct cggcctctct ggcggccttc tggcgctctt ccgcttcctc    1200

gctcactgac tcgctgcgct cggtcgttcg ctgcggcga gcggtatcag ctcactcaaa    1260

ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa    1320

aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct    1380

ccgcccccct gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac    1440

aggactataa agataccagg cgtttccccc tggaagctcc ctcgtgcgct ctcctgttcc    1500

gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttt    1560

ccgctgcata accctgcttc ggggtcatta tagcgatttt ttcggtatat ccatcctttt    1620
```

```
tcgcacgata tacaggattt tgccaaaggg ttcgtgtaga ctttccttgg tgtatccaac    1680

ggcgtcagcc gggcaggata ggtgaagtag gcccacccgc gagcgggtgt tccttcttca    1740

ctgtccctta ttcgcacctg gcggtgctca acgggaatcc tgctctgcga ggctggccgg    1800

ctaccgccgg cgtaacagat gagggcaagc ggatggctga tgaaaccaag ccaaccagga    1860

agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa    1920

aggcggcggc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc cagggctaca    1980

aaatcacggg cgtcgtggac tatgagcacg tccgcgagct ggcccgcatc aatggcgacc    2040

tgggccgcct gggcggcctg ctgaaactct ggctcaccga cgacccgcgc acggcgcggt    2100

tcggtgatgc cacgatcctc gccctgctgg cgaagatcga agagaagcag gacgagcttg    2160

gcaaggtcat gatgggcgtg gtccgcccga gggcagagcc atgacttttt tagccgctaa    2220

aacggccggg gggtgcgcgt gattgccaag cacgtcccca tgcgctccat caagaagagc    2280

gacttcgcgg agctggtgaa gtacatcacc gacgagcaag gcaagaccga gcgcctttgc    2340

gacgctcacc gggctggttg ccctcgccgc tgggctggcg gccgtctatg gccctgcaaa    2400

cgcgccagaa acgccgtcga agccgtgtgc gagacaccgc ggccgccggc gttgtggata    2460

cctcgcggaa aacttggccc tcactgacag atgaggggcg gacgttgaca cttgaggggc    2520

cgactcaccc ggcgcggcgt tgacagatga ggggcaggct cgatttcggc cggcgacgtg    2580

gagctggcca gcctcgcaaa tcggcgaaaa cgcctgattt tacgcgagtt cccacagat    2640

gatgtggaca agcctgggga taagtgccct gcggtattga cacttgaggg gcgcgactac    2700

tgacagatga ggggcgcgat ccttgacact tgaggggcag agtgctgaca gatgaggggc    2760

gcacctattg acatttgagg ggctgtccac aggcagaaaa tccagcattt gcaagggttt    2820

ccgcccgttt ttcggccacc gctaacctgt cttttaacct gcttttaaac caatatttat    2880

aaaccttgtt tttaaccagg ctgcgccct gtgcgcgtga ccgcgcacgc cgaaggggggg    2940

tgcccccct tctcgaaccc tcccggcccg ctaacgcggg cctcccatcc ccccaggggc    3000

tgcgcccctc ggccgcgaac ggcctcaccc caaaaatggc agcgctggca gtccttgcca    3060

ttgccgggat cggggcagta acgggatggg cgatcagccc gagcgcgacg cccggaagca    3120

ttgacgtgcc gcaggtgctg gcatcgacat tcagcgacca ggtgccgggc agtgagggcg    3180

gcggcctggg tggcggcctg cccttcactt cggccgtcgg ggcattcacg gacttcatgg    3240

cggggccggc aattttacc ttgggcattc ttggcatagt ggtcgcgggt gccgtgctcg    3300
```

```
tgttcggggg tgcgataaac ccagcgaacc atttgaggtg ataggtaaga ttataccgag    3360

gtatgaaaac gagaattgga cctttacaga attactctat gaagcgccat atttaaaaag    3420

ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgacaa    3480

tactgataag ataatatatc ttttatatag aagatatcgc cgtatgtaag gatttcaggg    3540

ggcaaggcat aggcagcgcg cttatcaata tatctataga atgggcaaag cataaaaact    3600

tgcatggact aatgcttgaa acccaggaca ataaccttat agcttgtaaa ttctatcata    3660

attgggtaat gactccaact tattgatagt gttttatgtt cagataatgc ccgatgactt    3720

tgtcatgcag ctccaccgat tttgagaacg acagcgactt ccgtcccagc cgtgccaggt    3780

gctgcctcag attcaggtta tgccgctcaa ttcgctgcgt atatcgcttg ctgattacgt    3840

gcagctttcc cttcaggcgg gattcataca gcggccagcc atccgtcatc catatcacca    3900

cgtcaaaggg tgacagcagg ctcataagac gccccagcgt cgccatagtg cgttcaccga    3960

atacgtgcgc aacaaccgtc ttccggagac tgtcatacgc gtaaaacagc cagcgctggc    4020

gcgatttagc cccgacatag ccccactgtt cgtccatttc cgcgcagacg atgacgtcac    4080

tgcccggctg tatgcgcgag gttaccgact gcggcctgag ttttttaagt gacgtaaaat    4140

cgtgttgagg ccaacgccca taatgcgggc tgttgcccgg catccaacgc cattcatggc    4200

catatcaatg attttctggt gcgtaccggg ttgagaagcg gtgtaagtga actgcagttg    4260

ccatgtttta cggcagtgag agcagagata gcgctgatgt ccggcggtgc ttttgccgtt    4320

acgcaccacc ccgtcagtag ctgaacagga gggacagctg atagacacag aagccactgg    4380

agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca cccataattg    4440

tggtttcaaa atcggctccg tcgatactat gttatacgcc aactttgaaa acaactttga    4500

aaaagctgtt ttctggtatt taaggtttta gaatgcaagg aacagtgaat tggagttcgt    4560

cttgttataa ttagcttctt ggggtatctt taaatactgt agaaagagg aaggaaataa     4620

taaatggcta aaatgagaat atcaccggaa ttgaaaaaac tgatcgaaaa ataccgctgc    4680

gtaaaagata cggaaggaat gtctcctgct aaggtatata agctggtggg agaaaatgaa    4740

aacctatatt taaaaatgac ggacagccgg tataaaggga ccacctatga tgtggaacgg    4800

gaaaaggaca tgatgctatg ctggaagga aagctgcctg ttccaaaggt cctgcacttt     4860

gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt cctttgctcg    4920

gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcatc    4980

aggctctttc actccatcga catatcggat tgtccctata cgaatagctt agacagccgc    5040
```

```
ttagccgaat tggattactt actgaataac gatctggccg atgtggattg cgaaaactgg    5100

gaagaagaca ctccatttaa agatccgcgc gagctgtatg attttttaaa gacggaaaag    5160

cccgaagagg aacttgtctt ttcccacggc gacctgggag acagcaacat ctttgtgaaa    5220

gatggcaaag taagtggctt tattgatctt gggagaagcg gcagggcgga caagtggtat    5280

gacattgcct tctgcgtccg gtcgatcagg gaggatatcg gggaagaaca gtatgtcgag    5340

ctattttttg acttactggg gatcaagcct gattgggaga aaataaaata ttatatttta    5400

ctggatgaat tgttttagta cctagatgtg gcgcaacgat gccggcgaca agcaggagcg    5460

caccgacttc ttccgcatca agtgttttgg ctctcaggcc gaggcccacg gcaagtattt    5520

gggcaagggg tcgctggtat tcgtgcaggg caagattcgg aataccaagt acgagaagga    5580

cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc tggacaccaa    5640

ggcaccaggc gggtcaaatc aggaataagg gcacattgcc ccggcgtgag tcggggcaat    5700

cccgcaagga gggtgaatga atcggacgtt tgaccggaag gcatacaggc aagaactgat    5760

cgacgcgggg ttttccgccg aggatgccga aaccatcgca agccgcaccg tcatgcgtgc    5820

gccccgcgaa accttccagt ccgtcggctc gatggtccag caagctacgg ccaagatcga    5880

gcgcgacagc gtgcaactgg ctccccctgc cctgcccgcg ccatcggccg ccgtggagcg    5940

ttcgcgtcgt ctcgaacagg aggcggcagg tttggcgaag tcgatgacca tcgacacgcg    6000

aggaactatg acgaccaaga agcgaaaaac cgccggcgag gacctggcaa aacaggtcag    6060

cgaggccaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg aaatgcagct    6120

ttccttgttc gatattgcgc cgtggccgga cacgatgcga gcgatgccaa acgacacggc    6180

ccgctctgcc ctgttcacca cgcgcaacaa gaaaatcccg cgcgaggcgc tgcaaaacaa    6240

ggtcattttc cacgtcaaca aggacgtgaa gatcacctac accggcgtcg agctgcgggc    6300

cgacgatgac gaactggtgt ggcagcaggt gttggagtac gcgaagcgca ccccctatcgg    6360

cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggt cgatcaatgg    6420

ccggtattac acgaaggccg aggaatgcct gtcgcgccta caggcgacgg cgatgggctt    6480

cacgtccgac cgcgttgggc acctggaatc ggtgtcgctg ctgcaccgct tccgcgtcct    6540

ggaccgtggc aagaaaacgt cccgttgcca ggtcctgatc gacgaggaaa tcgtcgtgct    6600

gtttgctggc gaccactaca cgaaattcat atgggagaag taccgcaagc tgtcgccgac    6660

ggcccgacgg atgttcgact atttcagctc gcaccgggag ccgtacccgc tcaagctgga    6720
```

```
aaccttccgc ctcatgtgcg gatcggattc cacccgcgtg aagaagtggc gcgagcaggt    6780

cggcgaagcc tgcgaagagt tgcgaggcag cggcctggtg gaacacgcct gggtcaatga    6840

tgacctggtg cattgcaaac gctagggcct tgtggggtca gttccggctg ggggttcagc    6900

agccagcgct ttactggcat ttcaggaaca agcgggcact gctcgacgca cttgcttcgc    6960

tcagtatcgc tcgggacgca cggcgcgctc tacgaactgc cgataaacag aggattaaaa    7020

ttgacaattg tgattaaggc tcagattcga cggcttggag cggccgacgt gcaggatttc    7080

cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tgttcgggtc cgtttacgag    7140

cacgaggaga aaaagcccat ggaggcgttc gctgaacggt tgcgagatgc cgtggcattc    7200

ggcgcctaca tcgacggcga gatcattggg ctgtcggtct tcaaacagga ggacggcccc    7260

aaggacgctc acaaggcgca tctgtccggc gttttcgtgg agcccgaaca gcgaggccga    7320

ggggtcgccg gtatgctgct gcgggcgttg ccggcgggtt tattgctcgt gatgatcgtc    7380

cgacagattc caacgggaat ctggtggatg cgcatcttca tcctcggcgc acttaatatt    7440

tcgctattct ggagcttgtt gtttatttcg gtctaccgcc tgccgggcgg ggtcgcggcg    7500

acggtaggcg ctgtgcagcc gctgatggtc gtgttcatct ctgccgctct gctaggtagc    7560

ccgatacgat tgatggcggt cctgggggct atttgcggaa ctgcgggcgt ggcgctgttg    7620

gtgttgacac caaacgcagc gctagatcct gtcggcgtcg cagcgggcct ggcgggggcg    7680

gtttccatgg cgttcggaac cgtgctgacc cgcaagtggc aacctcccgt gcctctgctc    7740

acctttaccg cctggcaact ggcggccgga ggacttctgc tcgttccagt agctttagtg    7800

tttgatccgc caatcccgat gcctacagga accaatgttc tcggcctggc gtggctcggc    7860

ctgatcggag cgggtttaac ctacttcctt tggttccggg ggatctcgcg actcgaacct    7920

acagttgttt ccttactggg ctttctcagc cccagatctg gggtcgatca gccggggatg    7980

catcaggccg acagtcggaa cttcgggtcc ccgacctgta ccattcggtg agcaatggat    8040

aggggagttg atatcgtcaa cgttcacttc taaagaaata gcgccactca gcttcctcag    8100

cggctttatc cagcgatttc ctattatgtc ggcatagttc tcaagatcga cagcctgtca    8160

cggttaagcg agaaatgaat aagaaggctg ataattcgga tctctgcgag ggagatgata    8220

tttgatcaca ggcagcaacg ctctgtcatc gttacaatca acatgctacc ctccgcgaga    8280

tcatccgtgt ttcaaacccg gcagcttagt tgccgttctt ccgaatagca tcggtaacat    8340

gagcaaagtc tgccgcctta caacggctct cccgctgacg ccgtcccgga ctgatgggct    8400

gcctgtatcg agtggtgatt ttgtgccgag ctgccggtcg gggagctgtt ggctggctgg    8460
```

```
tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg     8520

gacgttttta atgtactggg gtggtttttc ttttcaccag tgagacgggc aacagctgat     8580

tgcccttcac cgcctggccc tgagagagtt gcagcaagcg gtccacgctg gtttgcccca     8640

gcaggcgaaa atcctgtttg atggtggttc cgaaatcggc aaaatccctt ataaatcaaa     8700

agaatagccc gagatagggt tgagtgttgt tccagtttgg aacaagagtc cactattaaa     8760

gaacgtggac tccaacgtca aagggcgaaa aaccgtctat caggcgatg gcccactacg      8820

tgaaccatca cccaaatcaa gttttttggg gtcgaggtgc cgtaaagcac taaatcggaa     8880

ccctaaaggg agcccccgat ttagagcttg acggggaaag ccggcgaacg tggcgagaaa     8940

ggaagggaag aaagcgaaag gagcgggcgc cattcaggct gcgcaactgt tgggaagggc     9000

gatcggtgcg ggcctcttcg ctattacgcc agctggcgaa aggggatgt gctgcaaggc      9060

gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg     9120

aattaattcc catcttgaaa gaaatatagt ttaaatattt attgataaaa taacaagtca     9180

ggtattatag tccaagcaaa aacataaatt tattgatgca agtttaaatt cagaaatatt     9240

tcaataactg attatatcag ctggtacatt gccgtagatg aaagactgag tgcgatatta     9300

tgtgtaatac ataaattgat gatatagcta gcttagctca tcgggggatc cgtcgaagct     9360

agcttgggtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat gcgctgcgaa     9420

tcgggagcgg cgataccgta aagcacgagg aagcggtcag cccattcgcc gccaagctct     9480

tcagcaatat cacgggtagc caacgctatg tcctgatagc ggtccgccac acccagccgg     9540

ccacagtcga tgaatccaga aaagcggcca ttttccacca tgatattcgg caagcaggca     9600

tcgccatggg tcacgacgag atcctcgccg tcgggcatgc gcgccttgag cctggcgaac     9660

agttcggctg gcgcgagccc ctgatgctct tcgtccagat catcctgatc gacaagaccg     9720

gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg cttggtggtc gaatgggcag     9780

gtagccggat caagcgtatg cagccgccgc attgcatcag ccatgatgga tactttctcg     9840

gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgcccaa tagcagccag     9900

tcccttcccg cttcagtgac aacgtcgagc acagctgcgc aaggaacgcc cgtcgtggcc     9960

agccacgata gccgcgctgc ctcgtcctgc agttcattca gggcaccgga caggtcggtc    10020

ttgacaaaaa gaaccgggcg cccctgcgct gacagccgga acacggcggc atcagagcag    10080

ccgattgtct gttgtgccca gtcatagccg aatagcctct ccacccaagc ggccggagaa    10140
```

```
cctgcgtgca atccatcttg ttcaatccaa gctcccatgg gccctcgact agagtcgaga   10200

tctggattga gagtgaatat gagactctaa ttggataccg aggggaattt atggaacgtc   10260

agtggagcat ttttgacaag aaatatttgc tagctgatag tgaccttagg cgacttttga   10320

acgcgcaata atggtttctg acgtatgtgc ttagctcatt aaactccaga aacccgcggc   10380

tgagtggctc cttcaacgtt gcggttctgt cagttccaaa cgtaaaacgg cttgtcccgc   10440

gtcatcggcg ggggtcataa cgtgactccc ttaattctcc gctcatgatc ttgatcccct   10500

gcgccatcag atccttggcg gcaagaaagc catccagttt actttgcagg gcttcccaac   10560

cttaccagag ggcgccccag ctggcaattc cggttcgctt gctgtccata aaaccgccca   10620

gtctagctat cgccatgtaa gcccactgca agctacctgc tttctctttg cgcttgcgtt   10680

ttcccttgtc cagatagccc agtagctgac attcatccgg ggtcagcacc gtttctgcgg   10740

actggctttc tacgtgttcc gcttccttta gcagcccttg cgccctgagt gcttgcggca   10800

gcgtgaagct tgcatgcctg caggtcgacg gcgcgccgag ctcctcgagc aaatttacac   10860

attgccacta aacgtctaaa cccttgtaat ttgtttttgt tttactatgt gtgttatgta   10920

tttgatttgc gataaatttt tatatttggt actaaattta taacaccttt tatgctaacg   10980

tttgccaaca cttagcaatt tgcaagttga ttaattgatt ctaaattatt tttgtcttct   11040

aaatacatat actaatcaac tggaaatgta aatatttgct aatatttcta ctataggaga   11100

attaaagtga gtgaatatgg taccacaagg tttggagatt taattgttgc aatgctgcat   11160

ggatggcata tacaccaaac attcaataat tcttgaggat aataatggta ccacacaaga   11220

tttgaggtgc atgaacgtca cgtggacaaa aggtttagta attttcaag acaacaatgt    11280

taccacacac aagttttgag gtgcatgcat ggatgccctg tggaaagttt aaaaatattt   11340

tggaaatgat ttgcatggaa gccatgtgta aaaccatgac atccacttgg aggatgcaat   11400

aatgaagaaa actacaaatt tacatgcaac tagttatgca tgtagtctat ataatgagga   11460

ttttgcaata ctttcattca tacacactca ctaagtttta cacgattata atttcttcat   11520

agccagccca ccgcggtggg cggccgcctg cagtctagaa ggcctcctgc tttaatgaga   11580

tatgcgagac gcctatgatc gcatgatatt tgctttcaat tctgttgtgc acgttgtaaa   11640

aaacctgagc atgtgtagct cagatcctta ccgccggttt cggttcattc taatgaatat   11700

atcacccgtt actatcgtat ttttatgaat aatattctcc gttcaattta ctgattgtcc   11760

gtcgacgaat tcgagctcgg cgcgcctcta gaggatcgat gaattcagat cggctgagtg   11820

gctccttcaa cgttgcggtt ctgtcagttc caaacgtaaa acggcttgtc ccgcgtcatc   11880
```

```
ggcgggggtc ataacgtgac tcccttaatt ctccgctcat gatcagattg tcgtttcccg    11940

ccttcagttt aaactatcag tgtttgacag gatatattgg cgggtaaacc taagagaaaa    12000

gagcgtttat tagaataatc ggatatttaa aagggcgtga aaaggtttat ccttcgtcca    12060

tttgtatgtg catgccaacc acagggttcc cca                                 12093
```

<210> 28
<211> 12085
<212> DNA
<213> artificial sequence

<220>
<221> misc_feature
<222> (1)..(12085)
<223> plant expression vector with one promoter-terminator expression cassette

<400> 28
```
gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg atccgacagc      60

gcgcccagca caggtgcgca ggcaaattgc accaacgcat acagcgccag cagaatgcca     120

tagtgggcgg tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccgg cagcaccggc     180

ataatcaggc cgatgccgac agcgtcgagc gcgacagtgc tcagaattac gatcagggt     240

atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg aacgcgcgga     300

ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa gtgtcaagca     360

tgacaaagtt gcagccgaat acagtgatcc gtgccgccct ggacctgttg aacgaggtcg     420

gcgtagacgg tctgacgaca cgcaaactgg cggaacggtt ggggggttcag cagccggcgc     480

tttactggca cttcaggaac aagcgggcgc tgctcgacgc actggccgaa gccatgctgg     540

cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca tttctgatcg     600

ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg cgcatccatg     660

ccggcacgcg accgggcgca ccgcagatgg aaacggccga cgcgcagctt cgcttcctct     720

gcgaggcggg ttttтcggcc ggggacgccg tcaatgcgct gatgacaatc agctacttca     780

ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag cgcggcggca     840

ccgttgaaca ggctccgctc tcgccgctgt tgcgggccgc gatagacgcc ttcgacgaag     900

ccggtccgga cgcagcgttc gagcagggac tcgcggtgat tgtcgatgga ttggcgaaaa     960

ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc    1020
```

```
tgccggagcg caacccactc actacagcag agccatgtag acaacatccc ctcccccttt    1080

ccaccgcgtc agacgcccgt agcagcccgc tacgggcttt ttcatgccct gccctagcgt    1140

ccaagcctca cggccgcgct cggcctctct ggcggccttc tggcgctctt ccgcttcctc    1200

gctcactgac tcgctgcgct cggtcgttcg gctgcggcga gcggtatcag ctcactcaaa    1260

ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa    1320

aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct    1380

ccgcccccct gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac    1440

aggactataa agataccagg cgtttccccc tggaagctcc ctcgtgcgct ctcctgttcc    1500

gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttt    1560

ccgctgcata accctgcttc ggggtcatta tagcgatttt ttcggtatat ccatcctttt    1620

tcgcacgata tacaggattt tgccaagggg ttcgtgtaga ctttccttgg tgtatccaac    1680

ggcgtcagcc gggcaggata ggtgaagtag gcccacccgc gagcgggtgt ccttcttca     1740

ctgtccctta ttcgcacctg gcggtgctca acgggaatcc tgctctgcga ggctggccgg    1800

ctaccgccgg cgtaacagat gagggcaagc ggatggctga tgaaaccaag ccaaccagga    1860

agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa    1920

aggcggcggc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc cagggctaca    1980

aaatcacggg cgtcgtggac tatgagcacg tccgcgagct ggcccgcatc aatggcgacc    2040

tgggccgcct gggcggcctg ctgaaactct ggctcaccga cgacccgcgc acggcgcggt    2100

tcggtgatgc cacgatcctc gccctgctgg cgaagatcga agagaagcag gacgagcttg    2160

gcaaggtcat gatgggcgtg gtccgcccga gggcagagcc atgacttttt tagccgctaa    2220

aacggccggg gggtgcgcgt gattgccaag cacgtcccca tgcgctccat caagaagagc    2280

gacttcgcgg agctggtgaa gtacatcacc gacgagcaag gcaagaccga gcgcctttgc    2340

gacgctcacc gggctggttg ccctcgccgc tgggctggcg ccgtctatg ccctgcaaa     2400

cgcgccagaa acgccgtcga agccgtgtgc gagacaccgc ggccgccggc gttgtggata    2460

cctcgcggaa aacttggccc tcactgacag atgaggggcg gacgttgaca cttgaggggc    2520

cgactcaccc ggcgcggcgt tgacagatga ggggcaggct cgatttcggc cggcgacgtg    2580

gagctggcca gcctcgcaaa tcggcgaaaa cgcctgattt tacgcgagtt tcccacagat    2640

gatgtggaca agcctgggga taagtgccct gcggtattga cacttgaggg gcgcgactac    2700

tgacagatga ggggcgcgat ccttgacact tgaggggcag agtgctgaca gatgaggggc    2760
```

```
gcacctattg acatttgagg ggctgtccac aggcagaaaa tccagcattt gcaagggttt    2820

ccgcccgttt ttcggccacc gctaacctgt cttttaacct gcttttaaac caatatttat    2880

aaaccttgtt tttaaccagg gctgcgccct gtgcgcgtga ccgcgcacgc cgaaggggggg   2940

tgccccccct tctcgaaccc tcccggcccg ctaacgcggg cctcccatcc ccccaggggc    3000

tgcgcccctc ggccgcgaac ggcctcaccc caaaaatggc agcgctggca gtccttgcca    3060

ttgccgggat cggggcagta acgggatggg cgatcagccc gagcgcgacg cccggaagca    3120

ttgacgtgcc gcaggtgctg gcatcgacat tcagcgacca ggtgccgggc agtgagggcg    3180

gcggcctggg tggcggcctg cccttcactt cggccgtcgg ggcattcacg gacttcatgg    3240

cggggccggc aattttttacc ttgggcattc ttggcatagt ggtcgcgggt gccgtgctcg    3300

tgttcggggg tgcgataaac ccagcgaacc atttgaggtg ataggtaaga ttataccgag    3360

gtatgaaaac gagaattgga cctttacaga attactctat gaagcgccat atttaaaaag    3420

ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgacaa    3480

tactgataag ataatatatc ttttatatag aagatatcgc cgtatgtaag gatttcaggg    3540

ggcaaggcat aggcagcgcg cttatcaata tatctataga atgggcaaag cataaaaact    3600

tgcatggact aatgcttgaa acccaggaca ataaccttat agcttgtaaa ttctatcata    3660

attgggtaat gactccaact tattgatagt gttttatgtt cagataatgc ccgatgactt    3720

tgtcatgcag ctccaccgat tttgagaacg acagcgactt ccgtcccagc cgtgccaggt    3780

gctgcctcag attcaggtta tgccgctcaa ttcgctgcgt atatcgcttg ctgattacgt    3840

gcagctttcc cttcaggcgg gattcataca gcggccagcc atccgtcatc catatcacca    3900

cgtcaaaggg tgacagcagg ctcataagac gccccagcgt cgccatagtg cgttcaccga    3960

atacgtgcgc aacaaccgtc ttccggagac tgtcatacgc gtaaaacagc cagcgctggc    4020

gcgatttagc cccgacatag ccccactgtt cgtccatttc cgcgcagacg atgacgtcac    4080

tgcccggctg tatgcgcgag gttaccgact gcggcctgag ttttttaagt gacgtaaaat    4140

cgtgttgagg ccaacgccca taatgcgggc tgttgcccgg catccaacgc cattcatggc    4200

catatcaatg attttctggt gcgtaccggg ttgagaagcg gtgtaagtga actgcagttg    4260

ccatgtttta cggcagtgag agcagagata gcgctgatgt ccggcggtgc ttttgccgtt    4320

acgcaccacc ccgtcagtag ctgaacagga gggacagctg atagacacag aagccactgg    4380

agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca cccataattg    4440
```

```
tggtttcaaa atcggctccg tcgatactat gttatacgcc aactttgaaa acaactttga    4500

aaaagctgtt ttctggtatt taaggtttta gaatgcaagg aacagtgaat tggagttcgt    4560

cttgttataa ttagcttctt ggggtatctt taaatactgt agaaagagg aaggaaataa    4620

taaatggcta aaatgagaat atcaccggaa ttgaaaaaac tgatcgaaaa ataccgctgc    4680

gtaaaagata cggaaggaat gtctcctgct aaggtatata agctggtggg agaaaatgaa    4740

aacctatatt taaaaatgac ggacagccgg tataaaggga ccacctatga tgtggaacgg    4800

gaaaaggaca tgatgctatg gctggaagga aagctgcctg ttccaaaggt cctgcacttt    4860

gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt cctttgctcg    4920

gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcatc    4980

aggctctttc actccatcga catatcggat tgtccctata cgaatagctt agacagccgc    5040

ttagccgaat tggattactt actgaataac gatctggccg atgtggattg cgaaaactgg    5100

gaagaagaca ctccatttaa agatccgcgc gagctgtatg attttttaaa gacggaaaag    5160

cccgaagagg aacttgtctt ttcccacggc gacctgggag acagcaacat ctttgtgaaa    5220

gatggcaaag taagtggctt tattgatctt gggagaagcg gcagggcgga caagtggtat    5280

gacattgcct tctgcgtccg gtcgatcagg gaggatatcg gggaagaaca gtatgtcgag    5340

ctattttttg acttactggg gatcaagcct gattgggaga aataaaata ttatatttta    5400

ctggatgaat tgttttagta cctagatgtg gcgcaacgat gccggcgaca agcaggagcg    5460

caccgacttc ttccgcatca agtgttttgg ctctcaggcc gaggcccacg gcaagtattt    5520

gggcaagggg tcgctggtat tcgtgcaggg caagattcgg aataccaagt acgagaagga    5580

cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc tggacaccaa    5640

ggcaccaggc gggtcaaatc aggaataagg gcacattgcc ccggcgtgag tcggggcaat    5700

cccgcaagga gggtgaatga atcggacgtt tgaccggaag gcatacaggc aagaactgat    5760

cgacgcgggg ttttccgccg aggatgccga aaccatcgca agccgcaccg tcatgcgtgc    5820

gccccgcgaa accttccagt ccgtcggctc gatggtccag caagctacgg ccaagatcga    5880

gcgcgacagc gtgcaactgg ctccccctgc cctgcccgcg ccatcggccg ccgtggagcg    5940

ttcgcgtcgt ctcgaacagg aggcggcagg tttggcgaag tcgatgacca tcgacacgcg    6000

aggaactatg acgaccaaga agcgaaaaac cgccggcgag gacctggcaa aacaggtcag    6060

cgaggccaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg aaatgcagct    6120

ttccttgttc gatattgcgc cgtggccgga cacgatgcga gcgatgccaa acgacacggc    6180
```

```
ccgctctgcc ctgttcacca cgcgcaacaa gaaaatcccg cgcgaggcgc tgcaaaacaa      6240

ggtcattttc cacgtcaaca aggacgtgaa gatcacctac accggcgtcg agctgcgggc      6300

cgacgatgac gaactggtgt ggcagcaggt gttggagtac gcgaagcgca cccctatcgg      6360

cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggt cgatcaatgg      6420

ccggtattac acgaaggccg aggaatgcct gtcgcgccta caggcgacgg cgatgggctt      6480

cacgtccgac cgcgttgggc acctggaatc ggtgtcgctg ctgcaccgct tccgcgtcct      6540

ggaccgtggc aagaaaacgt cccgttgcca ggtcctgatc gacgaggaaa tcgtcgtgct      6600

gtttgctggc gaccactaca cgaaattcat atgggagaag taccgcaagc tgtcgccgac      6660

ggcccgacgg atgttcgact atttcagctc gcaccgggag ccgtacccgc tcaagctgga      6720

aaccttccgc ctcatgtgcg gatcggattc cacccgcgtg aagaagtggc gcgagcaggt      6780

cggcgaagcc tgcgaagagt tgcgaggcag cggcctggtg gaacacgcct gggtcaatga      6840

tgacctggtg cattgcaaac gctagggcct tgtggggtca gttccggctg ggggttcagc      6900

agccagcgct ttactggcat ttcaggaaca agcgggcact gctcgacgca cttgcttcgc      6960

tcagtatcgc tcgggacgca cggcgcgctc tacgaactgc cgataaacag aggattaaaa      7020

ttgacaattg tgattaaggc tcagattcga cggcttggag cggccgacgt gcaggatttc      7080

cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tgttcgggtc cgtttacgag      7140

cacgaggaga aaaagcccat ggaggcgttc gctgaacggt tgcgagatgc cgtggcattc      7200

ggcgcctaca tcgacggcga gatcattggg ctgtcggtct caaacagga ggacggcccc      7260

aaggacgctc acaaggcgca tctgtccggc gttttcgtgg agcccgaaca gcgaggccga      7320

ggggtcgccg gtatgctgct gcgggcgttg ccggcgggtt tattgctcgt gatgatcgtc      7380

cgacagattc caacgggaat ctggtggatg cgcatcttca tcctcggcgc acttaatatt      7440

tcgctattct ggagcttgtt gtttatttcg gtctaccgcc tgccgggcgg ggtcgcggcg      7500

acggtaggcg ctgtgcagcc gctgatggtc gtgttcatct ctgccgctct gctaggtagc      7560

ccgatacgat tgatggcggt cctgggggct atttgcggaa ctgcgggcgt ggcgctgttg      7620

gtgttgacac caaacgcagc gctagatcct gtcggcgtcg cagcgggcct ggcggggggcg      7680

gtttccatgg cgttcggaac cgtgctgacc cgcaagtggc aacctcccgt gcctctgctc      7740

acctttaccg cctggcaact ggcggccgga ggacttctgc tcgttccagt agctttagtg      7800

tttgatccgc caatcccgat gcctacagga accaatgttc tcggcctggc gtggctcggc      7860
```

```
ctgatcggag cgggtttaac ctacttcctt tggttccggg ggatctcgcg actcgaacct    7920

acagttgttt ccttactggg ctttctcagc cccagatctg gggtcgatca gccggggatg    7980

catcaggccg acagtcggaa cttcgggtcc ccgacctgta ccattcggtg agcaatggat    8040

aggggagttg atatcgtcaa cgttcacttc taaagaaata gcgccactca gcttcctcag    8100

cggctttatc cagcgatttc ctattatgtc ggcatagttc tcaagatcga cagcctgtca    8160

cggttaagcg agaaatgaat aagaaggctg ataattcgga tctctgcgag ggagatgata    8220

tttgatcaca ggcagcaacg ctctgtcatc gttacaatca acatgctacc ctccgcgaga    8280

tcatccgtgt ttcaaacccg gcagcttagt tgccgttctt ccgaatagca tcggtaacat    8340

gagcaaagtc tgccgcctta caacggctct cccgctgacg ccgtcccgga ctgatgggct    8400

gcctgtatcg agtggtgatt ttgtgccgag ctgccggtcg gggagctgtt ggctggctgg    8460

tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg    8520

gacgttttta atgtactggg gtggtttttc ttttcaccag tgagacgggc aacagctgat    8580

tgcccttcac cgcctggccc tgagagagtt gcagcaagcg gtccacgctg gtttgcccca    8640

gcaggcgaaa atcctgtttg atggtggttc cgaaatcggc aaaatccctt ataaatcaaa    8700

agaatagccc gagatagggt tgagtgttgt tccagtttgg aacaagagtc cactattaaa    8760

gaacgtggac tccaacgtca aagggcgaaa aaccgtctat cagggcgatg gcccactacg    8820

tgaaccatca cccaaatcaa gttttttggg gtcgaggtgc cgtaaagcac taaatcggaa    8880

ccctaaaggg agcccccgat ttagagcttg acggggaaag ccggcgaacg tggcgagaaa    8940

ggaagggaag aaagcgaaag gagcgggcgc cattcaggct gcgcaactgt tgggaagggc    9000

gatcggtgcg ggcctcttcg ctattacgcc agctggcgaa aggggggatgt gctgcaaggc    9060

gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg    9120

aattaattcc catcttgaaa gaaatatagt ttaaatattt attgataaaa taacaagtca    9180

ggtattatag tccaagcaaa aacataaatt tattgatgca agtttaaatt cagaaatatt    9240

tcaataactg attatatcag ctggtacatt gccgtagatg aaagactgag tgcgatatta    9300

tgtgtaatac ataaattgat gatatagcta gcttagctca tcgggggatc cgtcgaagct    9360

agcttgggtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat gcgctgcgaa    9420

tcgggagcgg cgataccgta aagcacgagg aagcggtcag cccattcgcc gccaagctct    9480

tcagcaatat cacgggtagc caacgctatg tcctgatagc ggtccgccac acccagccgg    9540

ccacagtcga tgaatccaga aaagcggcca ttttccacca tgatattcgg caagcaggca    9600
```

```
tcgccatggg tcacgacgag atcctcgccg tcgggcatgc gcgccttgag cctggcgaac   9660

agttcggctg gcgcgagccc ctgatgctct tcgtccagat catcctgatc gacaagaccg   9720

gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg cttggtggtc gaatgggcag   9780

gtagccggat caagcgtatg cagccgccgc attgcatcag ccatgatgga tactttctcg   9840

gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgcccaa tagcagccag   9900

tcccttcccg cttcagtgac aacgtcgagc acagctgcgc aaggaacgcc cgtcgtggcc   9960

agccacgata gccgcgctgc ctcgtcctgc agttcattca gggcaccgga caggtcggtc   10020

ttgacaaaaa gaaccgggcg cccctgcgct gacagccgga acacggcggc atcagagcag   10080

ccgattgtct gttgtgccca gtcatagccg aatagcctct ccacccaagc ggccggagaa   10140

cctgcgtgca atccatcttg ttcaatccaa gctcccatgg gccctcgact agagtcgaga   10200

tctggattga gagtgaatat gagactctaa ttggataccg aggggaattt atggaacgtc   10260

agtggagcat ttttgacaag aaatatttgc tagctgatag tgaccttagg cgacttttga   10320

acgcgcaata atggtttctg acgtatgtgc ttagctcatt aaactccaga aacccgcggc   10380

tgagtggctc cttcaacgtt gcggttctgt cagttccaaa cgtaaaacgg cttgtcccgc   10440

gtcatcggcg ggggtcataa cgtgactccc ttaattctcc gctcatgatc ttgatcccct   10500

gcgccatcag atccttggcg gcaagaaagc catccagttt actttgcagg gcttcccaac   10560

cttaccagag ggcgccccag ctggcaattc cggttcgctt gctgtccata aaaccgccca   10620

gtctagctat cgccatgtaa gcccactgca agctacctgc tttctctttg cgcttgcgtt   10680

ttcccttgtc cagatagccc agtagctgac attcatccgg ggtcagcacc gtttctgcgg   10740

actggctttc tacgtgttcc gcttccttta gcagcccttg cgccctgagt gcttgcggca   10800

gcgtgaagct tgcatgcctg caggtcgacg gcgcgccgag ctcctcgagc aaatttacac   10860

attgccacta aacgtctaaa cccttgtaat ttgtttttgt tttactatgt gtgttatgta   10920

tttgatttgc gataaatttt tatatttggt actaaattta taacaccttt tatgctaacg   10980

tttgccaaca cttagcaatt tgcaagttga ttaattgatt ctaaattatt tttgtcttct   11040

aaatacatat actaatcaac tggaaatgta aatatttgct aatatttcta ctataggaga   11100

attaaagtga gtgaatatgg taccacaagg tttggagatt taattgttgc aatgctgcat   11160

ggatggcata tacaccaaac attcaataat tcttgaggat aataatggta ccacacaaga   11220

tttgaggtgc atgaacgtca cgtggacaaa aggtttagta attttttcaag acaacaatgt  11280
```

```
taccacacac aagttttgag gtgcatgcat ggatgccctg tggaaagttt aaaaatattt    11340

tggaaatgat ttgcatggaa gccatgtgta aaaccatgac atccacttgg aggatgcaat    11400

aatgaagaaa actacaaatt tacatgcaac tagttatgca tgtagtctat ataatgagga    11460

ttttgcaata ctttcattca tacacactca ctaagtttta cacgattata atttcttcat    11520

agccagcgga tccgatatcg ggcccgctag cgttaaccct gctttaatga gatatgcgag    11580

acgcctatga tcgcatgata tttgctttca attctgttgt gcacgttgta aaaaacctga    11640

gcatgtgtag ctcagatcct taccgccggt ttcggttcat tctaatgaat atatcacccg    11700

ttactatcgt atttttatga ataatattct ccgttcaatt tactgattgt ccgtcgacga    11760

attcgagctc ggcgcgcctc tagaggatcg atgaattcag atcggctgag tggctccttc    11820

aacgttgcgg ttctgtcagt tccaaacgta aaacggcttg tcccgcgtca tcggcggggg    11880

tcataacgtg actcccttaa ttctccgctc atgatcagat tgtcgtttcc cgccttcagt    11940

ttaaactatc agtgtttgac aggatatatt ggcgggtaaa cctaagagaa aagagcgttt    12000

attagaataa tcggatattt aaaagggcgt gaaaaggttt atccttcgtc catttgtatg    12060

tgcatgccaa ccacagggtt cccca                                         12085
```

```
<210>  29
<211>  12079
<212>  DNA
<213>  artificial sequence

<220>
<221>  misc_feature
<222>  (1)..(12079)
<223>  plant expression vector with one promoter-terminator expression cassette

<400>  29
gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg atccgacagc    60

gcgcccagca caggtgcgca ggcaaattgc accaacgcat acagcgccag cagaatgcca    120

tagtgggcgg tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccgg cagcaccggc    180

ataatcaggc cgatgccgac agcgtcgagc gcgacagtgc tcagaattac gatcaggggt    240

atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg aacgcgcgga    300

ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa gtgtcaagca    360

tgacaaagtt gcagccgaat acagtgatcc gtgccgccct ggacctgttg aacgaggtcg    420

gcgtagacgg tctgacgaca cgcaaactgg cggaacggtt gggggttcag cagccggcgc    480

tttactggca cttcaggaac aagcgggcgc tgctcgacgc actggccgaa gccatgctgg    540
```

cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca tttctgatcg   600

ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg cgcatccatg   660

ccggcacgcg accgggcgca ccgcagatgg aaacggccga cgcgcagctt cgcttcctct   720

gcgaggcggg tttttcggcc ggggacgccg tcaatgcgct gatgacaatc agctacttca   780

ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag cgcggcggca   840

ccgttgaaca ggctccgctc tcgccgctgt tgcgggccgc gatagacgcc ttcgacgaag   900

ccggtccgga cgcagcgttc gagcagggac tcgcggtgat tgtcgatgga ttggcgaaaa   960

ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc   1020

tgccggagcg caacccactc actacagcag agccatgtag acaacatccc ctcccccttt   1080

ccaccgcgtc agacgcccgt agcagcccgc tacgggcttt ttcatgccct gccctagcgt   1140

ccaagcctca cggccgcgct cggcctctct ggcggccttc tggcgctctt ccgcttcctc   1200

gctcactgac tcgctgcgct cggtcgttcg gctgcggcga gcggtatcag ctcactcaaa   1260

ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa   1320

aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt ccataggct   1380

ccgcccccct gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac   1440

aggactataa agataccagg cgtttccccc tggaagctcc ctcgtgcgct ctcctgttcc   1500

gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttt   1560

ccgctgcata accctgcttc ggggtcatta tagcgatttt ttcggtatat ccatcctttt   1620

tcgcacgata tacaggattt tgccaaaggg ttcgtgtaga ctttccttgg tgtatccaac   1680

ggcgtcagcc gggcaggata ggtgaagtag gcccacccgc gagcgggtgt tccttcttca   1740

ctgtccctta ttcgcacctg gcggtgctca acgggaatcc tgctctgcga ggctggccgg   1800

ctaccgccgg cgtaacagat gagggcaagc ggatggctga tgaaaccaag ccaaccagga   1860

agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa   1920

aggcggcggc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc caggctaca   1980

aaatcacggg cgtcgtggac tatgagcacg tccgcgagct ggcccgcatc aatggcgacc   2040

tgggccgcct gggcggcctg ctgaaactct ggctcaccga cgacccgcgc acggcgcggt   2100

tcggtgatgc cacgatcctc gccctgctgg cgaagatcga agagaagcag gacgagcttg   2160

gcaaggtcat gatgggcgtg gtccgcccga gggcagagcc atgacttttt tagccgctaa   2220

```
aacggccggg gggtgcgcgt gattgccaag cacgtcccca tgcgctccat caagaagagc    2280

gacttcgcgg agctggtgaa gtacatcacc gacgagcaag gcaagaccga gcgcctttgc    2340

gacgctcacc gggctggttg ccctcgccgc tgggctggcg gccgtctatg gccctgcaaa    2400

cgcgccagaa acgccgtcga agccgtgtgc gagacaccgc ggccgccggc gttgtggata    2460

cctcgcggaa aacttggccc tcactgacag atgaggggcg gacgttgaca cttgaggggc    2520

cgactcaccc ggcgcggcgt tgacagatga ggggcaggct cgatttcggc cggcgacgtg    2580

gagctggcca gcctcgcaaa tcggcgaaaa cgcctgattt tacgcgagtt tcccacagat    2640

gatgtggaca agcctgggga taagtgccct gcggtattga cacttgaggg gcgcgactac    2700

tgacagatga ggggcgcgat ccttgacact tgaggggcag agtgctgaca gatgaggggc    2760

gcacctattg acatttgagg ggctgtccac aggcagaaaa tccagcattt gcaagggttt    2820

ccgcccgttt ttcggccacc gctaacctgt cttttaacct gcttttaaac caatatttat    2880

aaaccttgtt tttaaccagg gctgcgccct gtgcgcgtga ccgcgcacgc cgaaggggggg    2940

tgcccccccct tctcgaaccc tcccggcccg ctaacgcggg cctcccatcc ccccagggggc   3000

tgcgcccctc ggccgcgaac ggcctcaccc caaaaatggc agcgctggca gtccttgcca    3060

ttgccgggat cggggcagta acgggatggg cgatcagccc gagcgcgacg cccggaagca    3120

ttgacgtgcc gcaggtgctg gcatcgacat tcagcgacca ggtgccgggc agtgagggcg    3180

gcggcctggg tggcggcctg cccttcactt cggccgtcgg ggcattcacg gacttcatgg    3240

cggggccggc aatttttacc ttgggcattc ttggcatagt ggtcgcgggt gccgtgctcg    3300

tgttcggggg tgcgataaac ccagcgaacc atttgaggtg ataggtaaga ttataccgag    3360

gtatgaaaac gagaattgga cctttacaga attactctat gaagcgccat atttaaaaag    3420

ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgacaa    3480

tactgataag ataatatatc ttttatatag aagatatcgc cgtatgtaag gatttcaggg    3540


ggcaaggcat aggcagcgcg cttatcaata tatctataga atgggcaaag cataaaaact    3600

tgcatggact aatgcttgaa acccaggaca ataaccttat agcttgtaaa ttctatcata    3660

attgggtaat gactccaact tattgatagt gttttatgtt cagataatgc ccgatgactt    3720

tgtcatgcag ctccaccgat tttgagaacg acagcgactt ccgtcccagc cgtgccaggt    3780

gctgcctcag attcaggtta tgccgctcaa ttcgctgcgt atatcgcttg ctgattacgt    3840

gcagctttcc cttcaggcgg gattcataca gcggccagcc atccgtcatc catatcacca    3900
```

```
cgtcaaaggg  tgacagcagg  ctcataagac  gccccagcgt  cgccatagtg  cgttcaccga      3960

atacgtgcgc  aacaaccgtc  ttccggagac  tgtcatacgc  gtaaaacagc  cagcgctggc      4020

gcgatttagc  cccgacatag  ccccactgtt  cgtccatttc  cgcgcagacg  atgacgtcac      4080

tgcccggctg  tatgcgcgag  gttaccgact  gcggcctgag  tttttttaagt  gacgtaaaat     4140

cgtgttgagg  ccaacgccca  taatgcgggc  tgttgcccgg  catccaacgc  cattcatggc      4200

catatcaatg  attttctggt  gcgtaccggg  ttgagaagcg  gtgtaagtga  actgcagttg      4260

ccatgtttta  cggcagtgag  agcagagata  gcgctgatgt  ccggcggtgc  ttttgccgtt      4320

acgcaccacc  ccgtcagtag  ctgaacagga  gggacagctg  atagacacag  aagccactgg      4380

agcacctcaa  aaacaccatc  atacactaaa  tcagtaagtt  ggcagcatca  cccataattg      4440

tggtttcaaa  atcggctccg  tcgatactat  gttatacgcc  aactttgaaa  acaactttga      4500

aaaagctgtt  ttctggtatt  taaggtttta  gaatgcaagg  aacagtgaat  tggagttcgt      4560

cttgttataa  ttagcttctt  ggggtatctt  taaatactgt  agaaaagagg  aaggaaataa      4620

taaatggcta  aaatgagaat  atcaccggaa  ttgaaaaaac  tgatcgaaaa  ataccgctgc      4680

gtaaaagata  cggaaggaat  gtctcctgct  aaggtatata  agctggtggg  agaaaatgaa      4740

aacctatatt  taaaaatgac  ggacagccgg  tataaaggga  ccacctatga  tgtggaacgg      4800

gaaaaggaca  tgatgctatg  gctggaagga  aagctgcctg  ttccaaaggt  cctgcacttt      4860

gaacggcatg  atggctggag  caatctgctc  atgagtgagg  ccgatggcgt  cctttgctcg      4920

gaagagtatg  aagatgaaca  aagccctgaa  aagattatcg  agctgtatgc  ggagtgcatc      4980

aggctctttc  actccatcga  catatcggat  tgtccctata  cgaatagctt  agacagccgc      5040

ttagccgaat  tggattactt  actgaataac  gatctggccg  atgtggattg  cgaaaactgg      5100

gaagaagaca  ctccatttaa  agatccgcgc  gagctgtatg  attttttaaa  gacggaaaag      5160

cccgaagagg  aacttgtctt  ttcccacggc  gacctgggag  acagcaacat  ctttgtgaaa      5220

gatggcaaag  taagtggctt  tattgatctt  gggagaagcg  gcagggcgga  caagtggtat      5280

gacattgcct  tctgcgtccg  gtcgatcagg  gaggatatcg  gggaagaaca  gtatgtcgag      5340

ctattttttg  acttactggg  gatcaagcct  gattgggaga  aaataaaata  ttatattta      5400

ctggatgaat  tgtttttagta  cctagatgtg  gcgcaacgat  gccggcgaca  agcaggagcg      5460

caccgacttc  ttccgcatca  agtgttttgg  ctctcaggcc  gaggcccacg  gcaagtattt      5520

gggcaagggg  tcgctggtat  tcgtgcaggg  caagattcgg  aataccaagt  acgagaagga      5580

cggccagacg  gtctacggga  ccgacttcat  tgccgataag  gtggattatc  tggacaccaa      5640
```

```
ggcaccaggc gggtcaaatc aggaataagg gcacattgcc ccggcgtgag tcggggcaat      5700

cccgcaagga gggtgaatga atcggacgtt tgaccggaag gcatacaggc aagaactgat      5760

cgacgcgggg ttttccgccg aggatgccga aaccatcgca agccgcaccg tcatgcgtgc      5820

gccccgcgaa accttccagt ccgtcggctc gatggtccag caagctacgg ccaagatcga      5880

gcgcgacagc gtgcaactgg ctccccctgc cctgcccgcg ccatcggccg ccgtggagcg      5940

ttcgcgtcgt ctcgaacagg aggcggcagg tttggcgaag tcgatgacca tcgacacgcg      6000

aggaactatg acgaccaaga agcgaaaaac cgccggcgag gacctggcaa aacaggtcag      6060

cgaggccaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg aaatgcagct      6120

ttccttgttc gatattgcgc cgtggccgga cacgatgcga gcgatgccaa acgacacggc      6180

ccgctctgcc ctgttcacca cgcgcaacaa gaaaatcccg cgcgaggcgc tgcaaaacaa      6240

ggtcattttc cacgtcaaca aggacgtgaa gatcacctac accggcgtcg agctgcgggc      6300

cgacgatgac gaactggtgt ggcagcaggt gttggagtac gcgaagcgca cccctatcgg      6360

cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggt cgatcaatgg      6420

ccggtattac acgaaggccg aggaatgcct gtcgcgccta caggcgacgg cgatgggctt      6480

cacgtccgac cgcgttgggc acctggaatc ggtgtcgctg ctgcaccgct tccgcgtcct      6540

ggaccgtggc aagaaaacgt cccgttgcca ggtcctgatc gacgaggaaa tcgtcgtgct      6600

gtttgctggc gaccactaca cgaaattcat atgggagaag taccgcaagc tgtcgccgac      6660

ggcccgacgg atgttcgact atttcagctc gcaccgggag ccgtacccgc tcaagctgga      6720

aaccttccgc ctcatgtgcg gatcggattc cacccgcgtg aagaagtggc gcgagcaggt      6780

cggcgaagcc tgcgaagagt tgcgaggcag cggcctggtg gaacacgcct gggtcaatga      6840

tgacctggtg cattgcaaac gctagggcct tgtggggtca gttccggctg ggggttcagc      6900

agccagcgct ttactggcat ttcaggaaca agcgggcact gctcgacgca cttgcttcgc      6960

tcagtatcgc tcgggacgca cggcgcgctc tacgaactgc cgataaacag aggattaaaa      7020

ttgacaattg tgattaaggc tcagattcga cggcttggag cggccgacgt gcaggatttc      7080

cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tgttcgggtc cgtttacgag      7140

cacgaggaga aaaagcccat ggaggcgttc gctgaacggt tgcgagatgc cgtggcattc      7200

ggcgcctaca tcgacggcga gatcattggg ctgtcggtct tcaaacagga ggacggcccc      7260

aaggacgctc acaaggcgca tctgtccggc gttttcgtgg agcccgaaca gcgaggccga      7320
```

```
ggggtcgccg gtatgctgct gcgggcgttg ccggcgggtt tattgctcgt gatgatcgtc    7380

cgacagattc caacgggaat ctggtggatg cgcatcttca tcctcggcgc acttaatatt    7440

tcgctattct ggagcttgtt gtttatttcg gtctaccgcc tgccgggcgg ggtcgcggcg    7500

acggtaggcg ctgtgcagcc gctgatggtc gtgttcatct ctgccgctct gctaggtagc    7560

ccgatacgat tgatggcggt cctgggggct atttgcggaa ctgcgggcgt ggcgctgttg    7620

gtgttgacac caaacgcagc gctagatcct gtcggcgtcg cagcgggcct ggcgggggcg    7680

gtttccatgg cgttcggaac cgtgctgacc cgcaagtggc aacctcccgt gcctctgctc    7740

acctttaccg cctggcaact ggcggccgga ggacttctgc tcgttccagt agctttagtg    7800

tttgatccgc caatcccgat gcctacagga accaatgttc tcggcctggc gtggctcggc    7860

ctgatcggag cgggtttaac ctacttcctt tggttccggg ggatctcgcg actcgaacct    7920

acagttgttt ccttactggg ctttctcagc cccagatctg gggtcgatca gccggggatg    7980

catcaggccg acagtcggaa cttcgggtcc ccgacctgta ccattcggtg agcaatggat    8040

aggggagttg atatcgtcaa cgttcacttc taaagaaata gcgccactca gcttcctcag    8100

cggctttatc cagcgatttc ctattatgtc ggcatagttc tcaagatcga cagcctgtca    8160

cggttaagcg agaaatgaat aagaaggctg ataattcgga tctctgcgag ggagatgata    8220

tttgatcaca ggcagcaacg ctctgtcatc gttacaatca acatgctacc ctccgcgaga    8280

tcatccgtgt ttcaaacccg gcagcttagt tgccgttctt ccgaatagca tcggtaacat    8340

gagcaaagtc tgccgcctta caacggctct cccgctgacg ccgtcccgga ctgatgggct    8400

gcctgtatcg agtggtgatt ttgtgccgag ctgccggtcg gggagctgtt ggctggctgg    8460

tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg    8520

gacgttttta atgtactggg gtggtttttc ttttcaccag tgagacgggc aacagctgat    8580

tgcccttcac cgcctggccc tgagagagtt gcagcaagcg tccacgctg gtttgcccca     8640

gcaggcgaaa atcctgtttg atggtggttc cgaaatcggc aaaatccctt ataaatcaaa    8700

agaatagccc gagatagggt tgagtgttgt tccagtttgg aacaagagtc cactattaaa    8760

gaacgtggac tccaacgtca aagggcgaaa aaccgtctat caggcgatg gcccactacg     8820

tgaaccatca cccaaatcaa gttttttggg gtcgaggtgc cgtaaagcac taaatcggaa    8880

ccctaaaggg agcccccgat ttagagcttg acggggaaag ccggcgaacg tggcgagaaa    8940

ggaagggaag aaagcgaaag gagcgggcgc cattcaggct gcgcaactgt tgggaagggc    9000
```

```
gatcggtgcg ggcctcttcg ctattacgcc agctggcgaa aggggggatgt gctgcaaggc   9060

gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg   9120

aattaattcc catcttgaaa gaaatatagt ttaaatattt attgataaaa taacaagtca   9180

ggtattatag tccaagcaaa aacataaatt tattgatgca agtttaaatt cagaaatatt   9240

tcaataactg attatatcag ctggtacatt gccgtagatg aaagactgag tgcgatatta   9300

tgtgtaatac ataaattgat gatatagcta gcttagctca tcggggatc cgtcgaagct   9360

agcttgggtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat gcgctgcgaa   9420

tcgggagcgg cgataccgta aagcacgagg aagcggtcag cccattcgcc gccaagctct   9480

tcagcaatat cacgggtagc caacgctatg tcctgatagc ggtccgccac acccagccgg   9540

ccacagtcga tgaatccaga aaagcggcca ttttccacca tgatattcgg caagcaggca   9600

tcgccatggg tcacgacgag atcctcgccg tcgggcatgc gcgccttgag cctggcgaac   9660

agttcggctg gcgcgagccc ctgatgctct tcgtccagat catcctgatc gacaagaccg   9720

gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg cttggtggtc gaatgggcag   9780

gtagccggat caagcgtatg cagccgccgc attgcatcag ccatgatgga tactttctcg   9840

gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgcccaa tagcagccag   9900

tcccttcccg cttcagtgac aacgtcgagc acagctgcgc aaggaacgcc cgtcgtggcc   9960

agccacgata gccgcgctgc ctcgtcctgc agttcattca gggcaccgga caggtcggtc  10020

ttgacaaaaa gaaccgggcg cccctgcgct gacagccgga acacggcggc atcagagcag  10080

ccgattgtct gttgtgccca gtcatagccg aatagcctct ccacccaagc ggccggagaa  10140

cctgcgtgca atccatcttg ttcaatccaa gctcccatgg ccctcgact agagtcgaga  10200

tctggattga gagtgaatat gagactctaa ttggataccg aggggaattt atggaacgtc  10260

agtggagcat ttttgacaag aaatatttgc tagctgatag tgaccttagg cgacttttga  10320

acgcgcaata atggtttctg acgtatgtgc ttagctcatt aaactccaga aacccgcggc  10380

tgagtggctc cttcaacgtt gcggttctgt cagttccaaa cgtaaaacgg cttgtcccgc  10440

gtcatcggcg ggggtcataa cgtgactccc ttaattctcc gctcatgatc ttgatcccct  10500

gcgccatcag atccttggcg gcaagaaagc catccagttt actttgcagg gcttcccaac  10560

cttaccagag ggcgccccag ctggcaattc cggttcgctt gctgtccata aaaccgccca  10620

gtctagctat cgccatgtaa gcccactgca agctacctgc tttctctttg cgcttgcgtt  10680

ttcccttgtc cagatagccc agtagctgac attcatccgg ggtcagcacc gtttctgcgg  10740
```

```
actggctttc tacgtgttcc gcttccttta gcagcccttg cgccctgagt gcttgcggca   10800

gcgtgaagct tgcatgcctg caggtcgacg gcgcgccgag ctcctcgagc aaatttacac   10860

attgccacta aacgtctaaa cccttgtaat ttgtttttgt tttactatgt gtgttatgta   10920

tttgatttgc gataaatttt tatatttggt actaaattta taacaccttt tatgctaacg   10980

tttgccaaca cttagcaatt tgcaagttga ttaattgatt ctaaattatt tttgtcttct   11040

aaatacatat actaatcaac tggaaatgta aatatttgct aatatttcta ctataggaga   11100

attaaagtga gtgaatatgg taccacaagg tttggagatt taattgttgc aatgctgcat   11160

ggatggcata tacaccaaac attcaataat tcttgaggat aataatggta ccacacaaga   11220


tttgaggtgc atgaacgtca cgtggacaaa aggtttagta atttttcaag acaacaatgt   11280

taccacacac aagttttgag gtgcatgcat ggatgccctg tggaaagttt aaaaatattt   11340

tggaaatgat ttgcatggaa gccatgtgta aaaccatgac atccacttgg aggatgcaat   11400

aatgaagaaa actacaaatt tacatgcaac tagttatgca tgtagtctat ataatgagga   11460

ttttgcaata ctttcattca tacacactca ctaagtttta cacgattata atttcttcat   11520

agccagcaga tctgccggca tcgatcccgg gccatggcct gctttaatga gatatgcgag   11580

acgcctatga tcgcatgata tttgctttca attctgttgt gcacgttgta aaaaacctga   11640

gcatgtgtag ctcagatcct taccgccggt ttcggttcat tctaatgaat atatcacccg   11700

ttactatcgt atttttatga ataatattct ccgttcaatt tactgattgt ccgtcgacga   11760

gctcggcgcg cctctagagg atcgatgaat tcagatcggc tgagtggctc cttcaacgtt   11820

gcggttctgt cagttccaaa cgtaaaacgg cttgtcccgc gtcatcggcg ggggtcataa   11880

cgtgactccc ttaattctcc gctcatgatc agattgtcgt ttcccgcctt cagtttaaac   11940

tatcagtgtt tgacaggata tattggcggg taaacctaag agaaagagc gtttattaga   12000

ataatcggat atttaaaagg gcgtgaaaag gtttatcctt cgtccatttg tatgtgcatg   12060

ccaaccacag ggttccccca                                                12079
```

<223> plant expression vector with two promoter-terminator expression cassettes

<400> 30

```
gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg atccgacagc        60

gcgcccagca caggtgcgca ggcaaattgc accaacgcat acagcgccag cagaatgcca       120

tagtgggcgg tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccgg cagcaccggc       180

ataatcaggc cgatgccgac agcgtcgagc gcgacagtgc tcagaattac gatcaggggt       240

atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg aacgcgcgga       300

ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa gtgtcaagca       360

tgacaaagtt gcagccgaat acagtgatcc gtgccgccct ggacctgttg aacgaggtcg       420

gcgtagacgg tctgacgaca cgcaaactgg cggaacggtt gggggttcag cagccggcgc       480

tttactggca cttcaggaac aagcgggcgc tgctcgacgc actggccgaa gccatgctgg       540

cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca tttctgatcg       600

ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg cgcatccatg       660

ccggcacgcg accgggcgca ccgcagatgg aaacggccga cgcgcagctt cgcttcctct       720

gcgaggcggg tttttcggcc ggggacgccg tcaatgcgct gatgacaatc agctacttca       780

ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag cgcggcggca       840

ccgttgaaca ggctccgctc tcgccgctgt tgcgggccgc gatagacgcc ttcgacgaag       900

ccggtccgga cgcagcgttc gagcagggac tcgcggtgat tgtcgatgga ttggcgaaaa       960

ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc      1020

tgccggagcg caacccactc actacagcag agccatgtag acaacatccc ctcccccttt      1080

ccaccgcgtc agacgcccgt agcagcccgc tacgggcttt ttcatgccct gccctagcgt      1140

ccaagcctca cggccgcgct cggcctctct ggcggccttc tggcgctctt ccgcttcctc      1200

gctcactgac tcgctgcgct cggtcgttcg gctgcggcga gcggtatcag ctcactcaaa      1260

ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa      1320

aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt ccataggct       1380

ccgcccccct gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac      1440

aggactataa agataccagg cgtttccccc tggaagctcc ctcgtgcgct ctcctgttcc      1500

gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttt      1560

ccgctgcata accctgcttc ggggtcatta tagcgatttt ttcggtatat ccatcctttt      1620
```

```
tcgcacgata tacaggattt tgccaaaggg ttcgtgtaga cttttccttgg tgtatccaac    1680

ggcgtcagcc gggcaggata ggtgaagtag gcccacccgc gagcgggtgt tccttcttca    1740

ctgtccctta ttcgcacctg gcggtgctca acgggaatcc tgctctgcga ggctggccgg    1800

ctaccgccgg cgtaacagat gagggcaagc ggatggctga tgaaaccaag ccaaccagga    1860

agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa    1920

aggcggcggc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc cagggctaca    1980

aaatcacggg cgtcgtggac tatgagcacg tccgcgagct ggcccgcatc aatggcgacc    2040

tgggccgcct gggcggcctg ctgaaactct ggctcaccga cgacccgcgc acggcgcggt    2100

tcggtgatgc cacgatcctc gccctgctgg cgaagatcga agagaagcag gacgagcttg    2160

gcaaggtcat gatgggcgtg gtccgcccga gggcagagcc atgacttttt tagccgctaa    2220

aacggccggg gggtgcgcgt gattgccaag cacgtcccca tgcgctccat caagaagagc    2280

gacttcgcgg agctggtgaa gtacatcacc gacgagcaag gcaagaccga gcgcctttgc    2340

gacgctcacc gggctggttg ccctcgccgc tgggctggcg gccgtctatg gccctgcaaa    2400

cgcgccagaa acgccgtcga agccgtgtgc gagacaccgc ggccgccggc gttgtggata    2460

cctcgcggaa aacttggccc tcactgacag atgaggggcg gacgttgaca cttgaggggc    2520

cgactcaccc ggcgcggcgt tgacagatga ggggcaggct cgatttcggc cggcgacgtg    2580

gagctggcca gcctcgcaaa tcggcgaaaa cgcctgattt tacgcgagtt tcccacagat    2640

gatgtggaca agcctgggga taagtgccct gcggtattga cacttgaggg gcgcgactac    2700

tgacagatga ggggcgcgat ccttgacact tgaggggcag agtgctgaca gatgaggggc    2760

gcacctattg acatttgagg ggctgtccac aggcagaaaa tccagcattt gcaagggttt    2820

ccgcccgttt ttcggccacc gctaacctgt ctttttaacct gcttttaaac caatatttat    2880

aaaccttgtt tttaaccagg gctgcgccct gtgcgcgtga ccgcgcacgc cgaagggggg    2940

tgcccccct tctcgaaccc tcccggcccg ctaacgcggg cctcccatcc ccccaggggc    3000

tgcgcccctc ggccgcgaac ggcctcaccc caaaaatggc agcgctggca gtccttgcca    3060

ttgccgggat cggggcagta acgggatggg cgatcagccc gagcgcgacg cccggaagca    3120

ttgacgtgcc gcaggtgctg gcatcgacat tcagcgacca ggtgccgggc agtgagggcg    3180

gcggcctggg tggcggcctg cccttcactt cggccgtcgg ggcattcacg gacttcatgg    3240

cggggccggc aattttacc ttgggcattc ttggcatagt ggtcgcgggt gccgtgctcg    3300

tgttcggggg tgcgataaac ccagcgaacc atttgaggtg ataggtaaga ttataccgag    3360
```

```
gtatgaaaac gagaattgga cctttacaga attactctat gaagcgccat atttaaaaag   3420

ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgacaa   3480

tactgataag ataatatatc ttttatatag aagatatcgc cgtatgtaag gatttcaggg   3540

ggcaaggcat aggcagcgcg cttatcaata tatctataga atgggcaaag cataaaaact   3600

tgcatggact aatgcttgaa acccaggaca ataaccttat agcttgtaaa ttctatcata   3660

attgggtaat gactccaact tattgatagt gttttatgtt cagataatgc ccgatgactt   3720

tgtcatgcag ctccaccgat tttgagaacg acagcgactt ccgtcccagc cgtgccaggt   3780

gctgcctcag attcaggtta tgccgctcaa ttcgctgcgt atatcgcttg ctgattacgt   3840

gcagctttcc cttcaggcgg gattcataca gcggccagcc atccgtcatc catatcacca   3900

cgtcaaaggg tgacagcagg ctcataagac gccccagcgt cgccatagtg cgttcaccga   3960

atacgtgcgc aacaaccgtc ttccggagac tgtcatacgc gtaaaacagc cagcgctggc   4020

gcgatttagc cccgacatag ccccactgtt cgtccatttc cgcgcagacg atgacgtcac   4080

tgcccggctg tatgcgcgag gttaccgact gcggcctgag ttttttaagt gacgtaaaat   4140

cgtgttgagg ccaacgccca taatgcgggc tgttgcccgg catccaacgc cattcatggc   4200

catatcaatg attttctggt gcgtaccggg ttgagaagcg gtgtaagtga actgcagttg   4260

ccatgtttta cggcagtgag agcagagata gcgctgatgt ccggcggtgc ttttgccgtt   4320

acgcaccacc ccgtcagtag ctgaacagga gggacagctg atagacacag aagccactgg   4380

agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca cccataattg   4440

tggtttcaaa atcggctccg tcgatactat gttatacgcc aactttgaaa caactttga   4500

aaaagctgtt ttctggtatt taaggtttta gaatgcaagg aacagtgaat tggagttcgt   4560

cttgttataa ttagcttctt ggggtatctt taaatactgt agaaagagg aaggaaataa   4620

taaatggcta aaatgagaat atcaccggaa ttgaaaaaac tgatcgaaaa ataccgctgc   4680

gtaaaagata cggaaggaat gtctcctgct aaggtatata agctggtggg agaaaatgaa   4740

aacctatatt taaaaatgac ggacagccgg tataaaggga ccacctatga tgtggaacgg   4800

gaaaaggaca tgatgctatg gctggaagga agctgcctg ttccaaaggt cctgcacttt   4860

gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt cctttgctcg   4920

gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcatc   4980

aggctctttc actccatcga catatcggat tgtccctata cgaatagctt agacagccgc   5040
```

```
ttagccgaat  tggattactt  actgaataac  gatctggccg  atgtggattg  cgaaaactgg      5100

gaagaagaca  ctccatttaa  agatccgcgc  gagctgtatg  attttttaaa  gacggaaaag      5160

cccgaagagg  aacttgtctt  ttcccacggc  gacctgggag  acagcaacat  ctttgtgaaa      5220

gatggcaaag  taagtggctt  tattgatctt  gggagaagcg  gcagggcgga  caagtggtat      5280

gacattgcct  tctgcgtccg  gtcgatcagg  gaggatatcg  gggaagaaca  gtatgtcgag      5340

ctattttttg  acttactggg  gatcaagcct  gattgggaga  aaataaaata  ttatatttta      5400

ctggatgaat  tgttttagta  cctagatgtg  gcgcaacgat  gccggcgaca  agcaggagcg      5460

caccgacttc  ttccgcatca  agtgttttgg  ctctcaggcc  gaggcccacg  gcaagtattt      5520

gggcaagggg  tcgctggtat  tcgtgcaggg  caagattcgg  aataccaagt  acgagaagga      5580

cggccagacg  gtctacggga  ccgacttcat  tgccgataag  gtggattatc  tggacaccaa      5640

ggcaccaggc  gggtcaaatc  aggaataagg  gcacattgcc  ccggcgtgag  tcggggcaat      5700

cccgcaagga  gggtgaatga  atcggacgtt  tgaccggaag  gcatacaggc  aagaactgat      5760

cgacgcgggg  ttttccgccg  aggatgccga  aaccatcgca  agccgcaccg  tcatgcgtgc      5820

gccccgcgaa  accttccagt  ccgtcggctc  gatggtccag  caagctacgg  ccaagatcga      5880

gcgcgacagc  gtgcaactgg  ctccccctgc  cctgcccgcg  ccatcggccg  ccgtggagcg      5940

ttcgcgtcgt  ctcgaacagg  aggcggcagg  tttggcgaag  tcgatgacca  tcgacacgcg      6000

aggaactatg  acgaccaaga  agcgaaaaac  cgccggcgag  gacctggcaa  aacaggtcag      6060

cgaggccaag  caggccgcgt  tgctgaaaca  cacgaagcag  cagatcaagg  aaatgcagct      6120

ttccttgttc  gatattgcgc  cgtggccgga  cacgatgcga  gcgatgccaa  acgacacggc      6180

ccgctctgcc  ctgttcacca  cgcgcaacaa  gaaaatcccg  cgcgaggcgc  tgcaaaacaa      6240

ggtcattttc  cacgtcaaca  aggacgtgaa  gatcacctac  accggcgtcg  agctgcgggc      6300

cgacgatgac  gaactggtgt  ggcagcaggt  gttggagtac  gcgaagcgca  cccctatcgg      6360

cgagccgatc  accttcacgt  tctacgagct  ttgccaggac  ctgggctggt  cgatcaatgg      6420

ccggtattac  acgaaggccg  aggaatgcct  gtcgcgccta  caggcgacgg  cgatgggctt      6480

cacgtccgac  cgcgttgggc  acctggaatc  ggtgtcgctg  ctgcaccgct  tccgcgtcct      6540

ggaccgtggc  aagaaaacgt  cccgttgcca  ggtcctgatc  gacgaggaaa  tcgtcgtgct      6600

gtttgctggc  gaccactaca  cgaaattcat  atgggagaag  taccgcaagc  tgtcgccgac      6660

ggcccgacgg  atgttcgact  atttcagctc  gcaccgggag  ccgtacccgc  tcaagctgga      6720

aaccttccgc  ctcatgtgcg  gatcggattc  cacccgcgtg  aagaagtggc  gcgagcaggt      6780
```

```
cggcgaagcc tgcgaagagt tgcgaggcag cggcctggtg gaacacgcct gggtcaatga      6840

tgacctggtg cattgcaaac gctagggcct tgtggggtca gttccggctg ggggttcagc      6900

agccagcgct ttactggcat ttcaggaaca agcgggcact gctcgacgca cttgcttcgc      6960

tcagtatcgc tcgggacgca cggcgcgctc tacgaactgc cgataaacag aggattaaaa      7020

ttgacaattg tgattaaggc tcagattcga cggcttggag cggccgacgt gcaggatttc      7080

cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tgttcgggtc cgtttacgag      7140

cacgaggaga aaaagcccat ggaggcgttc gctgaacggt tgcgagatgc cgtggcattc      7200

ggcgcctaca tcgacggcga gatcattggg ctgtcggtct tcaaacagga ggacggcccc      7260

aaggacgctc acaaggcgca tctgtccggc gttttcgtgg agcccgaaca gcgaggccga      7320

ggggtcgccg gtatgctgct gcgggcgttg ccggcgggtt tattgctcgt gatgatcgtc      7380

cgacagattc caacgggaat ctggtggatg cgcatcttca tcctcggcgc acttaatatt      7440

tcgctattct ggagcttgtt gtttatttcg gtctaccgcc tgccgggcgg ggtcgcggcg      7500

acggtaggcg ctgtgcagcc gctgatggtc gtgttcatct ctgccgctct gctaggtagc      7560

ccgatacgat tgatggcggt cctggggggct atttgcggaa ctgcgggcgt ggcgctgttg      7620

gtgttgacac caaacgcagc gctagatcct gtcggcgtcg cagcgggcct ggcggggggcg      7680

gtttccatgg cgttcggaac cgtgctgacc cgcaagtggc aacctcccgt gcctctgctc      7740

acctttaccg cctggcaact ggcggccgga ggacttctgc tcgttccagt agctttagtg      7800

tttgatccgc caatcccgat gcctacagga accaatgttc tcggcctggc gtggctcggc      7860

ctgatcggag cgggtttaac ctacttcctt tggttccggg ggatctcgcg actcgaacct      7920

acagttgttt ccttactggg cttttctcagc cccagatctg gggtcgatca gccggggatg      7980

catcaggccg acagtcggaa cttcgggtcc ccgacctgta ccattcggtg agcaatggat      8040

aggggagttg atatcgtcaa cgttcacttc taaagaaata gcgccactca gcttcctcag      8100

cggctttatc cagcgatttc ctattatgtc ggcatagttc tcaagatcga cagcctgtca      8160

cggttaagcg agaaatgaat aagaaggctg ataattcgga tctctgcgag ggagatgata      8220

tttgatcaca ggcagcaacg ctctgtcatc gttacaatca acatgctacc ctccgcgaga      8280

tcatccgtgt ttcaaacccg gcagcttagt tgccgttctt ccgaatagca tcggtaacat      8340

gagcaaagtc tgccgcctta caacggctct cccgctgacg ccgtcccgga ctgatgggct      8400

gcctgtatcg agtggtgatt ttgtgccgag ctgccggtcg gggagctgtt ggctggctgg      8460
```

```
tggcaggata  tattgtggtg  taaacaaatt  gacgcttaga  caacttaata  acacattgcg   8520

gacgttttta  atgtactggg  gtggtttttc  ttttcaccag  tgagacgggc  aacagctgat   8580

tgcccttcac  cgcctggccc  tgagagagtt  gcagcaagcg  gtccacgctg  gtttgcccca   8640

gcaggcgaaa  atcctgtttg  atggtggttc  cgaaatcggc  aaaatccctt  ataaatcaaa   8700

agaatagccc  gagatagggt  tgagtgttgt  tccagtttgg  aacaagagtc  cactattaaa   8760

gaacgtggac  tccaacgtca  aagggcgaaa  aaccgtctat  cagggcgatg  gcccactacg   8820

tgaaccatca  cccaaatcaa  gttttttggg  gtcgaggtgc  cgtaaagcac  taaatcggaa   8880

ccctaaaggg  agcccccgat  ttagagcttg  acggggaaag  ccggcgaacg  tggcgagaaa   8940

ggaagggaag  aaagcgaaag  gagcgggcgc  cattcaggct  gcgcaactgt  tgggaagggc   9000

gatcggtgcg  ggcctcttcg  ctattacgcc  agctggcgaa  aggggatgt  gctgcaaggc   9060

gattaagttg  ggtaacgcca  gggttttccc  agtcacgacg  ttgtaaaacg  acggccagtg   9120

aattaattcc  catcttgaaa  gaaatatagt  ttaaatattt  attgataaaa  taacaagtca   9180

ggtattatag  tccaagcaaa  aacataaatt  tattgatgca  agtttaaatt  cagaaatatt   9240

tcaataactg  attatatcag  ctggtacatt  gccgtagatg  aaagactgag  tgcgatatta   9300

tgtgtaatac  ataaattgat  gatatagcta  gcttagctca  tcgggggatc  cgtcgaagct   9360

agcttgggtc  ccgctcagaa  gaactcgtca  agaaggcgat  agaaggcgat  gcgctgcgaa   9420

tcgggagcgg  cgataccgta  aagcacgagg  aagcggtcag  cccattcgcc  gccaagctct   9480

tcagcaatat  cacgggtagc  caacgctatg  tcctgatagc  ggtccgccac  acccagccgg   9540

ccacagtcga  tgaatccaga  aaagcggcca  ttttccacca  tgatattcgg  caagcaggca   9600

tcgccatggg  tcacgacgag  atcctcgccg  tcgggcatgc  gcgccttgag  cctggcgaac   9660

agttcggctg  gcgcgagccc  ctgatgctct  tcgtccagat  catcctgatc  gacaagaccg   9720

gcttccatcc  gagtacgtgc  tcgctcgatg  cgatgtttcg  cttggtggtc  gaatgggcag   9780

gtagccggat  caagcgtatg  cagccgccgc  attgcatcag  ccatgatgga  tactttctcg   9840

gcaggagcaa  ggtgagatga  caggagatcc  tgccccggca  cttcgcccaa  tagcagccag   9900

tcccttcccg  cttcagtgac  aacgtcgagc  acagctgcgc  aaggaacgcc  cgtcgtggcc   9960

agccacgata  gccgcgctgc  ctcgtcctgc  agttcattca  gggcaccgga  caggtcggtc  10020

ttgacaaaaa  gaaccgggcg  cccctgcgct  gacagccgga  acacggcggc  atcagagcag  10080

ccgattgtct  gttgtgccca  gtcatagccg  aatagcctct  ccacccaagc  ggccggagaa  10140

cctgcgtgca  atccatcttg  ttcaatccaa  gctcccatgg  gccctcgact  agagtcgaga  10200
```

```
tctggattga gagtgaatat gagactctaa ttggataccg aggggaattt atggaacgtc   10260

agtggagcat ttttgacaag aaatatttgc tagctgatag tgaccttagg cgacttttga   10320

acgcgcaata atggtttctg acgtatgtgc ttagctcatt aaactccaga aacccgcggc   10380

tgagtggctc cttcaacgtt gcggttctgt cagttccaaa cgtaaaacgg cttgtcccgc   10440

gtcatcggcg ggggtcataa cgtgactccc ttaattctcc gctcatgatc ttgatcccct   10500

gcgccatcag atccttggcg gcaagaaagc catccagttt actttgcagg gcttcccaac   10560

cttaccagag ggcgccccag ctggcaattc cggttcgctt gctgtccata aaaccgccca   10620

gtctagctat cgccatgtaa gcccactgca agctacctgc tttctctttg cgcttgcgtt   10680

ttcccttgtc cagatagccc agtagctgac attcatccgg ggtcagcacc gtttctgcgg   10740

actggctttc tacgtgttcc gcttccttta gcagcccttg cgccctgagt gcttgcggca   10800

gcgtgaagct tgcatgcctg caggtcgacg gcgcgccgag ctcctcgagc aaatttacac   10860

attgccacta aacgtctaaa cccttgtaat ttgtttttgt tttactatgt gtgttatgta   10920

tttgatttgc gataaatttt tatatttggt actaaattta taacaccttt tatgctaacg   10980

tttgccaaca cttagcaatt tgcaagttga ttaattgatt ctaaattatt tttgtcttct   11040

aaatacatat actaatcaac tggaaatgta aatatttgct aatatttcta ctataggaga   11100

attaaagtga gtgaatatgg taccacaagg tttggagatt taattgttgc aatgctgcat   11160

ggatggcata tacaccaaac attcaataat tcttgaggat aataatggta ccacacaaga   11220

tttgaggtgc atgaacgtca cgtggacaaa aggtttagta atttttcaag acaacaatgt   11280

taccacacac aagttttgag gtgcatgcat ggatgccctg tggaaagttt aaaaatattt   11340

tggaaatgat ttgcatggaa gccatgtgta aaaccatgac atccacttgg aggatgcaat   11400

aatgaagaaa actacaaatt tacatgcaac tagttatgca tgtagtctat ataatgagga   11460

ttttgcaata ctttcattca tacacactca ctaagtttta cacgattata atttcttcat   11520

agccagccca ccgcggtggg cggccgcctg cagtctagaa ggcctcctgc tttaatgaga   11580

tatgcgagac gcctatgatc gcatgatatt tgctttcaat tctgttgtgc acgttgtaaa   11640

aaacctgagc atgtgtagct cagatcctta ccgccggttt cggttcattc taatgaatat   11700

atcacccgtt actatcgtat ttttatgaat aatattctcc gttcaattta ctgattgtcc   11760

gtcgagcaaa tttacacatt gccactaaac gtctaaaccc ttgtaatttg tttttgtttt   11820

actatgtgtg ttatgtattt gatttgcgat aaattttat atttggtact aaatttataa   11880
```

```
caccttttat gctaacgttt gccaacactt agcaatttgc aagttgatta attgattcta   11940

aattattttt gtcttctaaa tacatatact aatcaactgg aaatgtaaat atttgctaat   12000

atttctacta taggagaatt aaagtgagtg aatatggtac cacaaggttt ggagatttaa   12060

ttgttgcaat gctgcatgga tggcatatac accaaacatt caataattct tgaggataat   12120

aatggtacca cacaagattt gaggtgcatg aacgtcacgt ggacaaaagg tttagtaatt   12180

tttcaagaca acaatgttac cacacacaag ttttgaggtg catgcatgga tgccctgtgg   12240

aaagtttaaa aatattttgg aaatgatttg catggaagcc atgtgtaaaa ccatgacatc   12300

cacttggagg atgcaataat gaagaaaact acaaatttac atgcaactag ttatgcatgt   12360

agtctatata atgaggattt tgcaatactt tcattcatac acactcacta agttttacac   12420

gattataatt tcttcatagc cagcggatcc gatatcgggc ccgctagcgt taaccctgct   12480

ttaatgagat atgcgagacg cctatgatcg catgatattt gctttcaatt ctgttgtgca   12540

cgttgtaaaa aacctgagca tgtgtagctc agatccttac cgccggtttc ggttcattct   12600

aatgaatata tcacccgtta ctatcgtatt tttatgaata atattctccg ttcaatttac   12660

tgattgtccg tcgacgaatt cgagctcggc gcgcctctag aggatcgatg aattcagatc   12720

ggctgagtgg ctccttcaac gttgcggttc tgtcagttcc aaacgtaaaa cggcttgtcc   12780

cgcgtcatcg gcgggggtca taacgtgact cccttaattc tccgctcatg atcagattgt   12840

cgtttcccgc cttcagttta aactatcagt gtttgacagg atatattggc gggtaaacct   12900

aagagaaaag agcgtttatt agaataatcg gatatttaaa agggcgtgaa aaggtttatc   12960

cttcgtccat ttgtatgtgc atgccaacca cagggttccc ca                      13002
```

```
<210>  31
<211>  13905
<212>  DNA
<213>  artificial sequence

<220>
<221>  misc_feature
<222>  (1)..(12085)
<223>  plant expression vector with three promoter-terminator expression
cassettes

<400>  31
gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg atccgacagc    60

gcgcccagca caggtgcgca ggcaaattgc accaacgcat acagcgccag cagaatgcca   120

tagtgggcgg tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccgg cagcaccggc   180
```

```
ataatcaggc cgatgccgac agcgtcgagc gcgacagtgc tcagaattac gatcaggggt    240

atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg aacgcgcgga    300

ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa gtgtcaagca    360


tgacaaagtt gcagccgaat acagtgatcc gtgccgccct ggacctgttg aacgaggtcg    420

gcgtagacgg tctgacgaca cgcaaactgg cggaacggtt gggggttcag cagccggcgc    480

tttactggca cttcaggaac aagcgggcgc tgctcgacgc actggccgaa gccatgctgg    540

cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca tttctgatcg    600

ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg cgcatccatg    660

ccggcacgcg accgggcgca ccgcagatgg aaacggccga cgcgcagctt cgcttcctct    720

gcgaggcggg tttttcggcc ggggacgccg tcaatgcgct gatgacaatc agctacttca    780

ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag cgcggcggca    840

ccgttgaaca ggctccgctc tcgccgctgt tgcgggccgc gatagacgcc ttcgacgaag    900

ccggtccgga cgcagcgttc gagcagggac tcgcggtgat tgtcgatgga ttggcgaaaa    960

ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc   1020

tgccggagcg caacccactc actacagcag agccatgtag acaacatccc ctcccccttt   1080

ccaccgcgtc agacgcccgt agcagcccgc tacgggcttt ttcatgccct gccctagcgt   1140

ccaagcctca cggccgcgct cggcctctct ggcggccttc tggcgctctt ccgcttcctc   1200

gctcactgac tcgctgcgct cggtcgttcg gctgcggcga gcggtatcag ctcactcaaa   1260

ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa   1320

aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt ccataggct   1380

ccgcccccct gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac   1440

aggactataa agataccagg cgtttccccc tggaagctcc ctcgtgcgct ctcctgttcc   1500

gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttt   1560

ccgctgcata accctgcttc ggggtcatta tagcgatttt ttcggtatat ccatcctttt   1620

tcgcacgata tacaggattt tgccaaaggg ttcgtgtaga ctttccttgg tgtatccaac   1680

ggcgtcagcc gggcaggata ggtgaagtag ccccacccgc gagcgggtgt tccttcttca   1740

ctgtccctta ttcgcacctg gcggtgctca acgggaatcc tgctctgcga ggctggccgg   1800

ctaccgccgg cgtaacagat gagggcaagc ggatggctga tgaaaccaag ccaaccagga   1860
```

```
agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa      1920

aggcggcggc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc cagggctaca      1980

aaatcacggg cgtcgtggac tatgagcacg tccgcgagct ggcccgcatc aatggcgacc      2040

tgggccgcct gggcggcctg ctgaaactct ggctcaccga cgacccgcgc acggcgcggt      2100

tcggtgatgc cacgatcctc gccctgctgg cgaagatcga agagaagcag gacgagcttg      2160

gcaaggtcat gatgggcgtg gtccgcccga gggcagagcc atgacttttt tagccgctaa      2220

aacggccggg gggtgcgcgt gattgccaag cacgtcccca tgcgctccat caagaagagc      2280

gacttcgcgg agctggtgaa gtacatcacc gacgagcaag gcaagaccga gcgcctttgc      2340

gacgctcacc gggctggttg ccctcgccgc tgggctggcg ccgtctatg gccctgcaaa      2400

cgcgccagaa acgccgtcga agccgtgtgc gagacaccgc ggccgccggc gttgtggata      2460

cctcgcggaa aacttggccc tcactgacag atgaggggcg gacgttgaca cttgaggggc      2520

cgactcaccc ggcgcggcgt tgacagatga ggggcaggct cgatttcggc cggcgacgtg      2580

gagctggcca gcctcgcaaa tcggcgaaaa cgcctgattt tacgcgagtt tcccacagat      2640

gatgtggaca agcctgggga taagtgccct gcggtattga' cacttgaggg gcgcgactac      2700

tgacagatga ggggcgcgat ccttgacact tgaggggcag agtgctgaca gatgaggggc      2760

gcacctattg acatttgagg ggctgtccac aggcagaaaa tccagcattt gcaagggttt      2820

ccgcccgttt ttcggccacc gctaacctgt cttttaacct gcttttaaac caatatttat      2880

aaaccttgtt tttaaccagg gctgcgccct gtgcgcgtga ccgcgcacgc cgaaggggg      2940

tgcccccccct tctcgaaccc tcccggcccg ctaacgcggg cctcccatcc ccccagggg      3000

tgcgccctc ggccgcgaac ggcctcaccc caaaaatggc agcgctggca gtccttgcca      3060

ttgccgggat cggggcagta acgggatggg cgatcagccc gagcgcgacg cccggaagca      3120

ttgacgtgcc gcaggtgctg gcatcgacat tcagcgacca ggtgccgggc agtgagggcg      3180

gcggcctggg tggcggcctg cccttcactt cggccgtcgg ggcattcacg gacttcatgg      3240

cggggccggc aattttacc ttgggcattc ttggcatagt ggtcgcgggt gccgtgctcg      3300

tgttcggggg tgcgataaac ccagcgaacc atttgaggtg ataggtaaga ttataccgag      3360

gtatgaaaac gagaattgga cctttacaga attactctat gaagcgccat atttaaaaag      3420

ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgacaa      3480

tactgataag ataatatatc ttttatatag aagatatcgc cgtatgtaag gatttcaggg      3540

ggcaaggcat aggcagcgcg cttatcaata tatctataga atgggcaaag cataaaaact      3600
```

```
tgcatggact aatgcttgaa acccaggaca ataaccttat agcttgtaaa ttctatcata    3660

attgggtaat gactccaact tattgatagt gttttatgtt cagataatgc ccgatgactt    3720

tgtcatgcag ctccaccgat tttgagaacg acagcgactt ccgtcccagc cgtgccaggt    3780

gctgcctcag attcaggtta tgccgctcaa ttcgctgcgt atatcgcttg ctgattacgt    3840

gcagctttcc cttcaggcgg gattcataca gcggccagcc atccgtcatc catatcacca    3900

cgtcaaaggg tgacagcagg ctcataagac gccccagcgt cgccatagtg cgttcaccga    3960

atacgtgcgc aacaaccgtc ttccggagac tgtcatacgc gtaaaacagc cagcgctggc    4020

gcgatttagc cccgacatag ccccactgtt cgtccatttc cgcgcagacg atgacgtcac    4080

tgcccggctg tatgcgcgag gttaccgact gcggcctgag ttttttaagt gacgtaaaat    4140

cgtgttgagg ccaacgccca taatgcgggc tgttgcccgg catccaacgc cattcatggc    4200


catatcaatg attttctggt gcgtaccggg ttgagaagcg gtgtaagtga actgcagttg    4260

ccatgtttta cggcagtgag agcagagata gcgctgatgt ccggcggtgc ttttgccgtt    4320

acgcaccacc ccgtcagtag ctgaacagga gggacagctg atagacacag aagccactgg    4380

agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca cccataattg    4440

tggtttcaaa atcggctccg tcgatactat gttatacgcc aactttgaaa acaactttga    4500

aaaagctgtt ttctggtatt taaggtttta gaatgcaagg aacagtgaat tggagttcgt    4560

cttgttataa ttagcttctt ggggtatctt taaatactgt agaaaagagg aaggaaataa    4620

taaatggcta aaatgagaat atcaccggaa ttgaaaaaac tgatcgaaaa ataccgctgc    4680

gtaaaagata cggaaggaat gtctcctgct aaggtatata agctggtggg agaaaatgaa    4740

aacctatatt taaaaatgac ggacagccgg tataaaggga ccacctatga tgtggaacgg    4800

gaaaaggaca tgatgctatg gctggaagga aagctgcctg ttccaaaggt cctgcacttt    4860

gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt cctttgctcg    4920

gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcatc    4980

aggctctttc actccatcga catatcggat tgtccctata cgaatagctt agacagccgc    5040

ttagccgaat tggattactt actgaataac gatctggccg atgtggattg cgaaaactgg    5100

gaagaagaca ctccatttaa agatccgcgc gagctgtatg attttttaaa gacggaaaag    5160

cccgaagagg aacttgtctt ttcccacggc gacctgggag acagcaacat ctttgtgaaa    5220

gatggcaaag taagtggctt tattgatctt gggagaagcg gcagggcgga caagtggtat    5280
```

```
gacattgcct tctgcgtccg gtcgatcagg gaggatatcg gggaagaaca gtatgtcgag    5340

ctattttttg acttactggg gatcaagcct gattgggaga aaataaaata ttatatttta    5400

ctggatgaat tgttttagta cctagatgtg gcgcaacgat gccggcgaca agcaggagcg    5460

caccgacttc ttccgcatca agtgttttgg ctctcaggcc gaggcccacg gcaagtattt    5520

gggcaagggg tcgctggtat tcgtgcaggg caagattcgg aataccaagt acgagaagga    5580

cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc tggacaccaa    5640

ggcaccaggc gggtcaaatc aggaataagg gcacattgcc ccggcgtgag tcggggcaat    5700

cccgcaagga gggtgaatga atcggacgtt tgaccggaag gcatacaggc aagaactgat    5760

cgacgcgggg ttttccgccg aggatgccga aaccatcgca agccgcaccg tcatgcgtgc    5820

gccccgcgaa accttccagt ccgtcggctc gatggtccag caagctacgg ccaagatcga    5880

gcgcgacagc gtgcaactgg ctccccctgc cctgcccgcg ccatcggccg ccgtggagcg    5940

ttcgcgtcgt ctcgaacagg aggcggcagg tttggcgaag tcgatgacca tcgacacgcg    6000

aggaactatg acgaccaaga agcgaaaaac cgccggcgag gacctggcaa aacaggtcag    6060

cgaggccaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg aaatgcagct    6120

ttccttgttc gatattgcgc cgtggccgga cacgatgcga gcgatgccaa acgacacggc    6180

ccgctctgcc ctgttcacca cgcgcaacaa gaaaatcccg cgcgaggcgc tgcaaaacaa    6240

ggtcattttc cacgtcaaca aggacgtgaa gatcacctac accggcgtcg agctgcgggc    6300

cgacgatgac gaactggtgt ggcagcaggt gttggagtac gcgaagcgca cccctatcgg    6360

cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggt cgatcaatgg    6420

ccggtattac acgaaggccg aggaatgcct gtcgcgccta caggcgacgg cgatgggctt    6480

cacgtccgac cgcgttgggc acctggaatc ggtgtcgctg ctgcaccgct tccgcgtcct    6540

ggaccgtggc aagaaaacgt cccgttgcca ggtcctgatc gacgaggaaa tcgtcgtgct    6600

gtttgctggc gaccactaca cgaaattcat atgggagaag taccgcaagc tgtcgccgac    6660

ggcccgacgg atgttcgact atttcagctc gcaccgggag ccgtacccgc tcaagctgga    6720

aaccttccgc ctcatgtgcg gatcggattc cacccgcgtg aagaagtggc gcgagcaggt    6780

cggcgaagcc tgcgaagagt tgcgaggcag cggcctggtg aacacgcct gggtcaatga     6840

tgacctggtg cattgcaaac gctagggcct tgtggggtca gttccggctg ggggttcagc    6900

agccagcgct ttactggcat ttcaggaaca agcgggcact gctcgacgca cttgcttcgc    6960
```

```
tcagtatcgc tcgggacgca cggcgcgctc tacgaactgc cgataaacag aggattaaaa   7020

ttgacaattg tgattaaggc tcagattcga cggcttggag cggccgacgt gcaggatttc   7080

cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tgttcgggtc cgtttacgag   7140

cacgaggaga aaaagcccat ggaggcgttc gctgaacggt tgcgagatgc cgtggcattc   7200

ggcgcctaca tcgacggcga gatcattggg ctgtcggtct tcaaacagga ggacggcccc   7260

aaggacgctc acaaggcgca tctgtccggc gttttcgtgg agcccgaaca gcgaggccga   7320

ggggtcgccg gtatgctgct gcgggcgttg ccggcgggtt tattgctcgt gatgatcgtc   7380

cgacagattc caacgggaat ctggtggatg cgcatcttca tcctcggcgc acttaatatt   7440

tcgctattct ggagcttgtt gtttatttcg gtctaccgcc tgccgggcgg ggtcgcggcg   7500

acggtaggcg ctgtgcagcc gctgatggtc gtgttcatct ctgccgctct gctaggtagc   7560

ccgatacgat tgatggcggt cctgggggct atttgcggaa ctgcgggcgt ggcgctgttg   7620

gtgttgacac caaacgcagc gctagatcct gtcggcgtcg cagcgggcct ggcgggggcg   7680

gtttccatgg cgttcggaac cgtgctgacc cgcaagtggc aacctcccgt gcctctgctc   7740

acctttaccg cctggcaact ggcggccgga ggacttctgc tcgttccagt agctttagtg   7800

tttgatccgc caatcccgat gcctacagga accaatgttc tcggcctggc gtggctcggc   7860

ctgatcggag cgggtttaac ctacttcctt tggttccggg ggatctcgcg actcgaacct   7920

acagttgttt ccttactggg ctttctcagc cccagatctg gggtcgatca gccggggatg   7980

catcaggccg acagtcggaa cttcgggtcc ccgacctgta ccattcggtg agcaatggat   8040


aggggagttg atatcgtcaa cgttcacttc taaagaaata gcgccactca gcttcctcag   8100

cggctttatc cagcgatttc ctattatgtc ggcatagttc tcaagatcga cagcctgtca   8160

cggttaagcg agaaatgaat aagaaggctg ataattcgga tctctgcgag ggagatgata   8220

tttgatcaca ggcagcaacg ctctgtcatc gttacaatca acatgctacc ctccgcgaga   8280

tcatccgtgt ttcaaacccg gcagcttagt tgccgttctt ccgaatagca tcggtaacat   8340

gagcaaagtc tgccgcctta caacggctct cccgctgacg ccgtcccgga ctgatgggct   8400

gcctgtatcg agtggtgatt ttgtgccgag ctgccggtcg gggagctgtt ggctggctgg   8460

tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg   8520

gacgttttta atgtactggg gtggtttttc ttttcaccag tgagacgggc aacagctgat   8580

tgcccttcac cgcctggccc tgagagagtt gcagcaagcg gtccacgctg gtttgcccca   8640
```

```
gcaggcgaaa atcctgtttg atggtggttc cgaaatcggc aaaatccctt ataaatcaaa   8700

agaatagccc gagatagggt tgagtgttgt tccagtttgg aacaagagtc cactattaaa   8760

gaacgtggac tccaacgtca aagggcgaaa aaccgtctat cagggcgatg gcccactacg   8820

tgaaccatca cccaaatcaa gttttttggg gtcgaggtgc cgtaaagcac taaatcggaa   8880

ccctaaaggg agcccccgat ttagagcttg acggggaaag ccggcgaacg tggcgagaaa   8940

ggaagggaag aaagcgaaag gagcgggcgc cattcaggct gcgcaactgt tgggaagggc   9000

gatcggtgcg ggcctcttcg ctattacgcc agctggcgaa agggggatgt gctgcaaggc   9060

gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg   9120

aattaattcc catcttgaaa gaaatatagt ttaaatattt attgataaaa taacaagtca   9180

ggtattatag tccaagcaaa aacataaatt tattgatgca agtttaaatt cagaaatatt   9240

tcaataactg attatatcag ctggtacatt gccgtagatg aaagactgag tgcgatatta   9300

tgtgtaatac ataaattgat gatatagcta gcttagctca tcgggggatc cgtcgaagct   9360

agcttgggtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat gcgctgcgaa   9420

tcgggagcgg cgataccgta aagcacgagg aagcggtcag cccattcgcc gccaagctct   9480

tcagcaatat cacgggtagc caacgctatg tcctgatagc ggtccgccac acccagccgg   9540

ccacagtcga tgaatccaga aaagcggcca ttttccacca tgatattcgg caagcaggca   9600

tcgccatggg tcacgacgag atcctcgccg tcgggcatgc gcgccttgag cctggcgaac   9660

agttcggctg gcgcgagccc ctgatgctct tcgtccagat catcctgatc gacaagaccg   9720

gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg cttggtggtc gaatgggcag   9780

gtagccggat caagcgtatg cagccgccgc attgcatcag ccatgatgga tactttctcg   9840

gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgcccaa tagcagccag   9900

tcccttcccg cttcagtgac aacgtcgagc acagctgcgc aaggaacgcc cgtcgtggcc  9960

agccacgata gccgcgctgc ctcgtcctgc agttcattca gggcaccgga caggtcggtc  10020

ttgacaaaaa gaaccgggcg cccctgcgct gacagccgga acacggcggc atcagagcag  10080

ccgattgtct gttgtgccca gtcatagccg aatagcctct ccacccaagc ggccggagaa  10140

cctgcgtgca atccatcttg ttcaatccaa gctcccatgg gccctcgact agagtcgaga  10200

tctggattga gagtgaatat gagactctaa ttggataccg aggggaattt atggaacgtc  10260

agtggagcat ttttgacaag aaatatttgc tagctgatag tgaccttagg cgacttttga  10320

acgcgcaata atggtttctg acgtatgtgc ttagctcatt aaactccaga aacccgcggc  10380
```

```
tgagtggctc cttcaacgtt gcggttctgt cagttccaaa cgtaaaacgg cttgtcccgc   10440

gtcatcggcg ggggtcataa cgtgactccc ttaattctcc gctcatgatc ttgatcccct   10500

gcgccatcag atccttggcg gcaagaaagc catccagttt actttgcagg gcttcccaac   10560

cttaccagag ggcgccccag ctggcaattc cggttcgctt gctgtccata aaaccgccca   10620

gtctagctat cgccatgtaa gcccactgca agctacctgc tttctctttg cgcttgcgtt   10680

ttcccttgtc cagatagccc agtagctgac attcatccgg ggtcagcacc gtttctgcgg   10740

actggctttc tacgtgttcc gcttccttta gcagcccttg cgccctgagt gcttgcggca   10800

gcgtgaagct tgcatgcctg caggtcgacg gcgcgccgag ctcctcgagc aaatttacac   10860

attgccacta aacgtctaaa cccttgtaat ttgtttttgt tttactatgt gtgttatgta   10920

tttgatttgc gataaatttt tatatttggt actaaattta taacaccttt tatgctaacg   10980

tttgccaaca cttagcaatt tgcaagttga ttaattgatt ctaaattatt tttgtcttct   11040

aaatacatat actaatcaac tggaaatgta aatatttgct aatatttcta ctataggaga   11100

attaaagtga gtgaatatgg taccacaagg tttggagatt taattgttgc aatgctgcat   11160

ggatggcata tacaccaaac attcaataat tcttgaggat aataatggta ccacacaaga   11220

tttgaggtgc atgaacgtca cgtggacaaa aggtttagta attttcaag acaacaatgt   11280

taccacacac aagttttgag gtgcatgcat ggatgccctg tggaaagttt aaaaatattt   11340

tggaaatgat ttgcatggaa gccatgtgta aaaccatgac atccacttgg aggatgcaat   11400

aatgaagaaa actacaaatt tacatgcaac tagttatgca tgtagtctat ataatgagga   11460

ttttgcaata ctttcattca tacacactca ctaagtttta cacgattata atttcttcat   11520

agccagccca ccgcggtggg cggccgcctg cagtctagaa ggcctcctgc tttaatgaga   11580

tatgcgagac gcctatgatc gcatgatatt tgctttcaat tctgttgtgc acgttgtaaa   11640

aaacctgagc atgtgtagct cagatcctta ccgccggttt cggttcattc taatgaatat   11700

atcacccgtt actatcgtat ttttatgaat aatattctcc gttcaattta ctgattgtcc   11760

gtcgagcaaa tttacacatt gccactaaac gtctaaaccc ttgtaatttg tttttgtttt   11820

actatgtgtg ttatgtattt gatttgcgat aaatttttat atttggtact aaatttataa   11880


cacctttat gctaacgttt gccaacactt agcaatttgc aagttgatta attgattcta   11940

aattattttt gtcttctaaa tacatatact aatcaactgg aaatgtaaat atttgctaat   12000

atttctacta taggagaatt aaagtgagtg aatatggtac cacaaggttt ggagatttaa   12060
```

```
ttgttgcaat gctgcatgga tggcatatac accaaacatt caataattct tgaggataat   12120

aatggtacca cacaagattt gaggtgcatg aacgtcacgt ggacaaaagg tttagtaatt   12180

tttcaagaca acaatgttac cacacacaag ttttgaggtg catgcatgga tgccctgtgg   12240

aaagtttaaa aatattttgg aaatgatttg catggaagcc atgtgtaaaa ccatgacatc   12300

cacttggagg atgcaataat gaagaaaact acaaatttac atgcaactag ttatgcatgt   12360

agtctatata atgaggattt tgcaatactt tcattcatac acactcacta agttttacac   12420

gattataatt tcttcatagc cagcggatcc gatatcgggc ccgctagcgt taaccctgct   12480

ttaatgagat atgcgagacg cctatgatcg catgatattt gctttcaatt ctgttgtgca   12540

cgttgtaaaa aacctgagca tgtgtagctc agatccttac cgccggtttc ggttcattct   12600

aatgaatata tcacccgtta ctatcgtatt tttatgaata atattctccg ttcaatttac   12660

tgattgtccg tcgagcaaat ttacacattg ccactaaacg tctaaaccct tgtaatttgt   12720

ttttgtttta ctatgtgtgt tatgtatttg atttgcgata aatttttata tttggtacta   12780

aatttataac accttttatg ctaacgtttg ccaacactta gcaatttgca agttgattaa   12840

ttgattctaa attatttttg tcttctaaat acatatacta atcaactgga aatgtaaata   12900

tttgctaata tttctactat aggagaatta aagtgagtga atatggtacc acaaggtttg   12960

gagatttaat tgttgcaatg ctgcatggat ggcatataca ccaaacattc aataattctt   13020

gaggataata atggtaccac acaagatttg aggtgcatga acgtcacgtg gacaaaaggt   13080

ttagtaattt ttcaagacaa caatgttacc acacacaagt tttgaggtgc atgcatggat   13140

gccctgtgga aagtttaaaa atattttgga aatgatttgc atggaagcca tgtgtaaaac   13200

catgacatcc acttggagga tgcaataatg aagaaaacta caaatttaca tgcaactagt   13260

tatgcatgta gtctatataa tgaggatttt gcaatacttt cattcataca cactcactaa   13320

gttttacacg attataattt cttcatagcc agcagatctg ccggcatcga tcccgggcca   13380

tggcctgctt taatgagata tgcgagacgc ctatgatcgc atgatatttg ctttcaattc   13440

tgttgtgcac gttgtaaaaa acctgagcat gtgtagctca gatccttacc gccggtttcg   13500

gttcattcta atgaatatat cacccgttac tatcgtattt ttatgaataa tattctccgt   13560

tcaatttact gattgtccgt cgacgagctc ggcgcgcctc tagaggatcg atgaattcag   13620

atcggctgag tggctccttc aacgttgcgg ttctgtcagt tccaaacgta aaacggcttg   13680

tcccgcgtca tcggcggggg tcataacgtg actcccttaa ttctccgctc atgatcagat   13740
```

```
tgtcgtttcc cgccttcagt ttaaactatc agtgtttgac aggatatatt ggcgggtaaa    13800

cctaagagaa aagagcgttt attagaataa tcggatattt aaaagggcgt gaaaaggttt    13860

atccttcgtc catttgtatg tgcatgccaa ccacagggtt cccca                    13905
```

```
<210>    32
<211>    1443
<212>    DNA
<213>    Phaeodactylum tricornutum

<220>
<221>    CDS
<222>    (9)..(1442)
<223>    delta-6-desaturase


<400>    32
gatctaaa atg ggc aaa gga ggg gac gct cgg gcc tcg aag ggc tca acg       50
         Met Gly Lys Gly Gly Asp Ala Arg Ala Ser Lys Gly Ser Thr
          1               5                      10

gcg gct cgc aag atc agt tgg cag gaa gtc aag acc cac gcg tct ccg       98
Ala Ala Arg Lys Ile Ser Trp Gln Glu Val Lys Thr His Ala Ser Pro
15                  20                  25                  30

gag gac gcc tgg atc att cac tcc aat aag gtc tac gac gtg tcc aac       146
Glu Asp Ala Trp Ile Ile His Ser Asn Lys Val Tyr Asp Val Ser Asn
                35                  40                  45

tgg cac gaa cat ccc gga ggc gcc gtc att ttc acg cac gcc ggt gac       194
Trp His Glu His Pro Gly Gly Ala Val Ile Phe Thr His Ala Gly Asp
                50                  55                  60

gac atg acg gac att ttc gct gcc ttt cac gca ccc gga tcg cag tcg       242
Asp Met Thr Asp Ile Phe Ala Ala Phe His Ala Pro Gly Ser Gln Ser
        65                  70                  75

ctc atg aag aag ttc tac att ggc gaa ttg ctc ccg gaa acc acc ggc       290
Leu Met Lys Lys Phe Tyr Ile Gly Glu Leu Leu Pro Glu Thr Thr Gly
        80                  85                  90

aag gag ccg cag caa atc gcc ttt gaa aag ggc tac cgc gat ctg cgc       338
Lys Glu Pro Gln Gln Ile Ala Phe Glu Lys Gly Tyr Arg Asp Leu Arg
95                  100                 105                 110

tcc aaa ctc atc atg atg ggc atg ttc aag tcc aac aag tgg ttc tac       386
Ser Lys Leu Ile Met Met Gly Met Phe Lys Ser Asn Lys Trp Phe Tyr
                115                 120                 125

gtc tac aag tgc ctc agc aac atg gcc att tgg gcc gcc gcc tgt gct       434
Val Tyr Lys Cys Leu Ser Asn Met Ala Ile Trp Ala Ala Ala Cys Ala
                130                 135                 140

ctc gtc ttt tac tcg gac cgc ttc tgg gta cac ctg gcc agc gcc gtc       482
Leu Val Phe Tyr Ser Asp Arg Phe Trp Val His Leu Ala Ser Ala Val
```

145   150   155

atg ctg gga aca ttc ttt cag cag tcg gga tgg ttg gca cac gac ttt   530
Met Leu Gly Thr Phe Phe Gln Gln Ser Gly Trp Leu Ala His Asp Phe
    160   165   170

ctg cac cac cag gtc ttc acc aag cgc aag cac ggg gat ctc gga gga   578
Leu His His Gln Val Phe Thr Lys Arg Lys His Gly Asp Leu Gly Gly
175   180   185   190

ctc ttt tgg ggg aac ctc atg cag ggt tac tcc gta cag tgg tgg aaa   626
Leu Phe Trp Gly Asn Leu Met Gln Gly Tyr Ser Val Gln Trp Trp Lys
                        195   200   205

aac aag cac aac gga cac cac gcc gtc ccc aac ctc cac tgc tcc tcc   674
Asn Lys His Asn Gly His His Ala Val Pro Asn Leu His Cys Ser Ser
                    210   215   220

gca gtc gcg caa gat ggg gac ccg gac atc gat acc atg ccc ctt ctc   722
Ala Val Ala Gln Asp Gly Asp Pro Asp Ile Asp Thr Met Pro Leu Leu
            225   230   235

gcc tgg tcc gtc cag caa gcc cag tct tac cgg gaa ctc caa gcc gac   770
Ala Trp Ser Val Gln Gln Ala Gln Ser Tyr Arg Glu Leu Gln Ala Asp
        240   245   250

gga aag gat tcg ggt ttg gtc aag ttc atg atc cgt aac caa tcc tac   818
Gly Lys Asp Ser Gly Leu Val Lys Phe Met Ile Arg Asn Gln Ser Tyr
255   260   265   270

ttt tac ttt ccc atc ttg ttg ctc gcc cgc ctg tcg tgg ttg aac gag   866
Phe Tyr Phe Pro Ile Leu Leu Leu Ala Arg Leu Ser Trp Leu Asn Glu
                    275   280   285

tcc ttc aag tgc gcc ttt ggg ctt gga gct gcg tcg gag aac gct gct   914
Ser Phe Lys Cys Ala Phe Gly Leu Gly Ala Ala Ser Glu Asn Ala Ala
                290   295   300

ctc gaa ctc aag gcc aag ggt ctt cag tac ccc ctt ttg gaa aag gct   962
Leu Glu Leu Lys Ala Lys Gly Leu Gln Tyr Pro Leu Leu Glu Lys Ala
            305   310   315

ggc atc ctg ctg cac tac gct tgg atg ctt aca gtt tcg tcc ggc ttt   1010
Gly Ile Leu Leu His Tyr Ala Trp Met Leu Thr Val Ser Ser Gly Phe
    320   325   330

gga cgc ttc tcg ttc gcg tac acc gca ttt tac ttt cta acc gcg acc   1058
Gly Arg Phe Ser Phe Ala Tyr Thr Ala Phe Tyr Phe Leu Thr Ala Thr
335   340   345   350

gcg tcc tgt gga ttc ttg ctc gcc att gtc ttt ggc ctc ggc cac aac   1106
Ala Ser Cys Gly Phe Leu Leu Ala Ile Val Phe Gly Leu Gly His Asn
                355   360   365

ggc atg gcc acc tac aat gcc gac gcc cgt ccg gac ttc tgg aag ctc   1154
Gly Met Ala Thr Tyr Asn Ala Asp Ala Arg Pro Asp Phe Trp Lys Leu
            370   375   380

```
caa gtc acc acg act cgc aac gtc acg ggc gga cac ggt ttc ccc caa       1202
Gln Val Thr Thr Thr Arg Asn Val Thr Gly Gly His Gly Phe Pro Gln
        385                 390             395

gcc ttt gtc gac tgg ttc tgt ggt ggc ctc cag tac caa gtc gac cac       1250
Ala Phe Val Asp Trp Phe Cys Gly Gly Leu Gln Tyr Gln Val Asp His
        400             405             410

cac tta ttc ccc agc ctg ccc cga cac aat ctg gcc aag aca cac gca       1298
His Leu Phe Pro Ser Leu Pro Arg His Asn Leu Ala Lys Thr His Ala
415             420             425             430

ctg gtc gaa tcg ttc tgc aag gag tgg ggt gtc cag tac cac gaa gcc       1346
Leu Val Glu Ser Phe Cys Lys Glu Trp Gly Val Gln Tyr His Glu Ala
                435             440             445

gac ctt gtg gac ggg acc atg gaa gtc ttg cac cat ttg ggc agc gtg       1394
Asp Leu Val Asp Gly Thr Met Glu Val Leu His His Leu Gly Ser Val
                450             455             460

gcc ggc gaa ttc gtc gtg gat ttt gta cgc gat gga ccc gcc atg taa a     1443
Ala Gly Glu Phe Val Val Asp Phe Val Arg Asp Gly Pro Ala Met
        465             470             475
```

<210> 33
<211> 477
<212> PRT
<213> Phaeodactylum tricornutum

<400> 33

```
Met Gly Lys Gly Gly Asp Ala Arg Ala Ser Lys Gly Ser Thr Ala Ala
1               5                   10                  15


Arg Lys Ile Ser Trp Gln Glu Val Lys Thr His Ala Ser Pro Glu Asp
            20                  25                  30


Ala Trp Ile Ile His Ser Asn Lys Val Tyr Asp Val Ser Asn Trp His
        35                  40                  45


Glu His Pro Gly Gly Ala Val Ile Phe Thr His Ala Gly Asp Asp Met
    50                  55                  60


Thr Asp Ile Phe Ala Ala Phe His Ala Pro Gly Ser Gln Ser Leu Met
65                  70                  75                  80


Lys Lys Phe Tyr Ile Gly Glu Leu Leu Pro Glu Thr Thr Gly Lys Glu
                85                  90                  95
```

Pro Gln Gln Ile Ala Phe Glu Lys Gly Tyr Arg Asp Leu Arg Ser Lys
        100   105   110

Leu Ile Met Met Gly Met Phe Lys Ser Asn Lys Trp Phe Tyr Val Tyr
        115   120   125

Lys Cys Leu Ser Asn Met Ala Ile Trp Ala Ala Ala Cys Ala Leu Val
    130   135   140

Phe Tyr Ser Asp Arg Phe Trp Val His Leu Ala Ser Ala Val Met Leu
145   150   155   160

Gly Thr Phe Phe Gln Gln Ser Gly Trp Leu Ala His Asp Phe Leu His
                165   170   175

His Gln Val Phe Thr Lys Arg Lys His Gly Asp Leu Gly Gly Leu Phe
                180   185   190

Trp Gly Asn Leu Met Gln Gly Tyr Ser Val Gln Trp Trp Lys Asn Lys
        195   200   205

His Asn Gly His His Ala Val Pro Asn Leu His Cys Ser Ser Ala Val
    210   215   220

Ala Gln Asp Gly Asp Pro Asp Ile Asp Thr Met Pro Leu Leu Ala Trp
225   230   235   240

Ser Val Gln Gln Ala Gln Ser Tyr Arg Glu Leu Gln Ala Asp Gly Lys
                245   250   255

Asp Ser Gly Leu Val Lys Phe Met Ile Arg Asn Gln Ser Tyr Phe Tyr
        260   265   270

Phe Pro Ile Leu Leu Leu Ala Arg Leu Ser Trp Leu Asn Glu Ser Phe
        275   280   285

Lys Cys Ala Phe Gly Leu Gly Ala Ala Ser Glu Asn Ala Ala Leu Glu
    290   295   300

Leu Lys Ala Lys Gly Leu Gln Tyr Pro Leu Leu Glu Lys Ala Gly Ile
305   310   315   320

Leu Leu His Tyr Ala Trp Met Leu Thr Val Ser Ser Gly Phe Gly Arg

                325   330   335

Phe Ser Phe Ala Tyr Thr Ala Phe Tyr Phe Leu Thr Ala Thr Ala Ser
            340   345   350

Cys Gly Phe Leu Leu Ala Ile Val Phe Gly Leu Gly His Asn Gly Met
            355   360   365

Ala Thr Tyr Asn Ala Asp Ala Arg Pro Asp Phe Trp Lys Leu Gln Val
    370   375   380

Thr Thr Thr Arg Asn Val Thr Gly Gly His Gly Phe Pro Gln Ala Phe
385   390   395   400

Val Asp Trp Phe Cys Gly Gly Leu Gln Tyr Gln Val Asp His His Leu
            405   410   415

Phe Pro Ser Leu Pro Arg His Asn Leu Ala Lys Thr His Ala Leu Val
            420   425   430

Glu Ser Phe Cys Lys Glu Trp Gly Val Gln Tyr His Glu Ala Asp Leu
            435   440   445

Val Asp Gly Thr Met Glu Val Leu His His Leu Gly Ser Val Ala Gly
    450   455   460

Glu Phe Val Val Asp Phe Val Arg Asp Gly Pro Ala Met
465   470   475

<210> 34
<211> 17061
<212> DNA
<213> Phaeodactylum tricornutum, Physcomitrella patens, Caenorhabditis elegans

<220>
<221> CDS
<222> (4554)..(5987)
<223> Phaeodactylum tricornutum delta-6-desaturase

<220>
<221> CDS
<222> (2805)..(3653)
<223> Caenorhabditis elegans LPLAT

<220>

```
<221>  CDS
<222>  (1026)..(1898)
<223>  Physcomitrella patens delta-6-elongase


<400>  34
tggggaaccc tgtggttggc atgcacatac aaatggacga aggataaacc ttttcacgcc        60

cttttaaata tccgattatt ctaataaacg ctctttctc ttaggtttac ccgccaatat        120

atcctgtcaa acactgatag tttaaactga aggcgggaaa cgacaatctg atcatgagcg        180

gagaattaag ggagtcacgt tatgaccccc gccgatgacg cgggacaagc cgttttacgt        240

ttggaactga cagaaccgca acgttgaagg agccactcag ccgatctgaa ttcatcgatc        300

ctctagaggc gcgccgagct cctcgagcaa atttacacat tgccactaaa cgtctaaacc        360

cttgtaattt gttttgtttt tactatgtgt gttatgtatt tgatttgcga taaattttta        420

tatttggtac taaatttata acacctttta tgctaacgtt tgccaacact tagcaatttg        480

caagttgatt aattgattct aaattatttt tgtcttctaa atacatatac taatcaactg        540

gaaatgtaaa tatttgctaa tatttctact ataggagaat taaagtgagt gaatatggta        600

ccacaaggtt tggagattta attgttgcaa tgctgcatgg atggcatata caccaaacat        660

tcaataattc ttgaggataa taatggtacc acacaagatt tgaggtgcat gaacgtcacg        720

tggacaaaag gtttagtaat ttttcaagac aacaatgtta ccacacacaa gttttgaggt        780

gcatgcatgg atgccctgtg gaaagtttaa aaatattttg gaaatgattt gcatggaagc        840

catgtgtaaa accatgacat ccacttggag gatgcaataa tgaagaaaac tacaaattta        900

catgcaacta gttatgcatg tagtctatat aatgaggatt ttgcaatact ttcattcata        960

cacactcact aagttttaca cgattataat ttcttcatag ccagcccacc gcggtgggcg       1020
```

```
gccgc atg gag gtc gtg gag aga ttc tac ggt gag ttg gat ggg aag gtc       1070
      Met Glu Val Val Glu Arg Phe Tyr Gly Glu Leu Asp Gly Lys Val
      1                   5                   10                  15

tcg cag ggc gtg aat gca ttg ctg ggt agt ttt ggg gtg gag ttg acg        1118
Ser Gln Gly Val Asn Ala Leu Leu Gly Ser Phe Gly Val Glu Leu Thr
                20              25                  30

gat acg ccc act acc aaa ggc ttg ccc ctc gtt gac agt ccc aca ccc        1166
Asp Thr Pro Thr Thr Lys Gly Leu Pro Leu Val Asp Ser Pro Thr Pro
            35                  40                  45

atc gtc ctc ggt gtt tct gta tac ttg act att gtc att gga ggg ctt        1214
Ile Val Leu Gly Val Ser Val Tyr Leu Thr Ile Val Ile Gly Gly Leu
        50                  55                  60

ttg tgg ata aag gcc agg gat ctg aaa ccg cgc gcc tcg gag cca ttt        1262
Leu Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg Ala Ser Glu Pro Phe
```

```
Leu Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg Ala Ser Glu Pro Phe
    65                  70              75
```

```
ttg ctc caa gct ttg gtg ctt gtg cac aac ctg ttc tgt ttt gcg ctc        1310
Leu Leu Gln Ala Leu Val Leu Val His Asn Leu Phe Cys Phe Ala Leu
80                  85                  90                  95
```

```
agt ctg tat atg tgc gtg ggc atc gct tat cag gct att acc tgg cgg        1358
Ser Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln Ala Ile Thr Trp Arg
                100                 105                 110
```

```
tac tct ctc tgg ggc aat gca tac aat cct aaa cat aaa gag atg gcg        1406
Tyr Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys His Lys Glu Met Ala
                115                 120                 125
```

```
att ctg gta tac ttg ttc tac atg tct aag tac gtg gaa ttc atg gat        1454
Ile Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr Val Glu Phe Met Asp
            130                 135                 140
```

```
acc gtt atc atg ata ctg aag cgc agc acc agg caa ata agc ttc ctc        1502
Thr Val Ile Met Ile Leu Lys Arg Ser Thr Arg Gln Ile Ser Phe Leu
        145                 150                 155
```

```
cac gtt tat cat cat tct tca att tcc ctc att tgg tgg gct att gct        1550
His Val Tyr His His Ser Ser Ile Ser Leu Ile Trp Trp Ala Ile Ala
160                 165                 170                 175
```

```
cat cac gct cct ggc ggt gaa gca tat tgg tct gcg gct ctg aac tca        1598
His His Ala Pro Gly Gly Glu Ala Tyr Trp Ser Ala Ala Leu Asn Ser
                180                 185                 190
```

```
gga gtg cat gtt ctc atg tat gcg tat tac ttc ttg gct gcc tgc ctt        1646
Gly Val His Val Leu Met Tyr Ala Tyr Tyr Phe Leu Ala Ala Cys Leu
                195                 200                 205
```

```
cga agt agc cca aag tta aaa aat aag tac ctt ttt tgg ggc agg tac        1694
Arg Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu Phe Trp Gly Arg Tyr
        210                 215                 220
```

```
ttg aca caa ttc caa atg ttc cag ttt atg ctg aac tta gtg cag gct        1742
Leu Thr Gln Phe Gln Met Phe Gln Phe Met Leu Asn Leu Val Gln Ala
        225                 230                 235
```

```
tac tac gac atg aaa acg aat gcg cca tat cca caa tgg ctg atc aag        1790
Tyr Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro Gln Trp Leu Ile Lys
240                 245                 250                 255
```

```
att ttg ttc tac tac atg atc tcg ttg ctg ttt ctt ttc ggc aat ttt        1838
Ile Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe Leu Phe Gly Asn Phe
                260                 265                 270
```

```
tac gta caa aaa tac atc aaa ccc tct gac gga aag caa aag gga gct        1886
Tyr Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly Lys Gln Lys Gly Ala
                275                 280                 285
```

```
aaa act gag tga tctagaaggc ctcctgcttt aatgagatat gcgagacgcc            1938
Lys Thr Glu
```

290

tatgatcgca tgatatttgc tttcaattct gttgtgcacg ttgtaaaaaa cctgagcatg   1998

tgtagctcag atccttaccg ccggtttcgg ttcattctaa tgaatatatc acccgttact   2058

atcgtatttt tatgaataat attctccgtt caatttactg attgtccgtc gagcaaattt   2118

acacattgcc actaaacgtc taaacccttg taatttgttt ttgttttact atgtgtgtta   2178

tgtatttgat ttgcgataaa tttttatatt tggtactaaa tttataacac cttttatgct   2238

aacgtttgcc aacacttagc aatttgcaag ttgattaatt gattctaaat tatttttgtc   2298

ttctaaatac atatactaat caactggaaa tgtaaatatt tgctaatatt tctactatag   2358

gagaattaaa gtgagtgaat atggtaccac aaggtttgga gatttaattg ttgcaatgct   2418

gcatggatgg catatacacc aaacattcaa taattcttga ggataataat ggtaccacac   2478

aagatttgag gtgcatgaac gtcacgtgga caaaaggttt agtaattttt caagacaaca   2538

atgttaccac acacaagttt tgaggtgcat gcatggatgc cctgtggaaa gtttaaaaat   2598

attttggaaa tgatttgcat ggaagccatg tgtaaaacca tgacatccac ttggaggatg   2658

caataatgaa gaaaactaca aatttacatg caactagtta tgcatgtagt ctatataatg   2718

aggattttgc aatactttca ttcatacaca ctcactaagt tttacacgat tataatttct   2778

tcatagccag cggatccgcc cacata atg gag aac ttc tgg tct att gtt gtg   2831
                                       Met Glu Asn Phe Trp Ser Ile Val Val
                                                  295

ttt ttt cta ctc tca att ctc ttc att tta tat aac ata tcg aca gta   2879
Phe Phe Leu Leu Ser Ile Leu Phe Ile Leu Tyr Asn Ile Ser Thr Val
300   305   310   315

tgc cac tac tat atg cgg att tcg ttt tat tac ttc aca att tta ttg   2927
Cys His Tyr Tyr Met Arg Ile Ser Phe Tyr Tyr Phe Thr Ile Leu Leu
              320   325   330

cat gga atg gaa gtt tgt gtt aca atg atc cct tct tgg cta aat ggg   2975
His Gly Met Glu Val Cys Val Thr Met Ile Pro Ser Trp Leu Asn Gly
             335   340   345

aag ggt gct gat tac gtg ttt cac tcg ttt ttc tat tgg tgt aaa tgg   3023
Lys Gly Ala Asp Tyr Val Phe His Ser Phe Phe Tyr Trp Cys Lys Trp
          350   355   360

act ggt gtt cat aca aca gtc tat gga tat gaa aaa aca caa gtt gaa   3071
Thr Gly Val His Thr Thr Val Tyr Gly Tyr Glu Lys Thr Gln Val Glu
          365   370   375

ggt ccg gct gta gtt att tgt aat cat cag agt tct ctc gac att cta   3119
Gly Pro Ala Val Val Ile Cys Asn His Gln Ser Ser Leu Asp Ile Leu
380   385   390   395

```
tcg atg gca tca atc tgg ccg aag aat tgt gtt gta atg atg aaa cga        3167
Ser Met Ala Ser Ile Trp Pro Lys Asn Cys Val Val Met Met Lys Arg
            400             405             410

att ctt gcc tat gtt cca ttc ttc aat ctc gga gcc tac ttt tcc aac        3215
Ile Leu Ala Tyr Val Pro Phe Phe Asn Leu Gly Ala Tyr Phe Ser Asn
            415             420             425

aca atc ttc atc gat cga tat aac cgt gaa cgt gcg atg gct tca gtt        3263
Thr Ile Phe Ile Asp Arg Tyr Asn Arg Glu Arg Ala Met Ala Ser Val
            430             435             440

gat tat tgt gca tct gaa atg aag aac aga aat ctt aaa ctt tgg gta        3311
Asp Tyr Cys Ala Ser Glu Met Lys Asn Arg Asn Leu Lys Leu Trp Val
            445             450             455

ttt ccg gaa gga aca aga aat cgt gaa gga ggg ttc att cca ttc aag        3359
Phe Pro Glu Gly Thr Arg Asn Arg Glu Gly Gly Phe Ile Pro Phe Lys
460             465             470             475

aaa gga gca ttc aat att gca gtt cgt gcg cag att ccc att att cca        3407
Lys Gly Ala Phe Asn Ile Ala Val Arg Ala Gln Ile Pro Ile Ile Pro
            480             485             490

gtt gta ttc tca gac tat cgg gat ttc tac tca aag cca ggc cga tat        3455
Val Val Phe Ser Asp Tyr Arg Asp Phe Tyr Ser Lys Pro Gly Arg Tyr
            495             500             505

ttc aag aat gat gga gaa gtt gtt att cga gtt ctg gat gcg att cca        3503
Phe Lys Asn Asp Gly Glu Val Val Ile Arg Val Leu Asp Ala Ile Pro
            510             515             520

aca aaa ggg ctc act ctt gat gac gtc agc gag ttg tct gat atg tgt        3551
Thr Lys Gly Leu Thr Leu Asp Asp Val Ser Glu Leu Ser Asp Met Cys
            525             530             535

cgg gac gtt atg ttg gca gcc tat aag gaa gtt act cta gaa gct cag        3599
Arg Asp Val Met Leu Ala Ala Tyr Lys Glu Val Thr Leu Glu Ala Gln
540             545             550             555

caa cga aat gcg aca cgg cgt gga gaa aca aaa gac ggg aag aaa tct        3647
Gln Arg Asn Ala Thr Arg Arg Gly Glu Thr Lys Asp Gly Lys Lys Ser
            560             565             570

gag taa gctagcgtta accctgcttt aatgagatat gcgagacgcc tatgatcgca        3703
Glu

tgatatttgc tttcaattct gttgtgcacg ttgtaaaaaa cctgagcatg tgtagctcag        3763

atccttaccg ccggtttcgg ttcattctaa tgaatatatc acccgttact atcgtatttt        3823

tatgaataat attctccgtt caatttactg attgtccgtc gagcaaattt acacattgcc        3883

actaaacgtc taaacccttg taatttgttt ttgtttttact atgtgtgtta tgtatttgat        3943
```

ttgcgataaa tttttatatt tggtactaaa tttataacac cttttatgct aacgtttgcc   4003

aacacttagc aatttgcaag ttgattaatt gattctaaat tatttttgtc ttctaaatac   4063

atatactaat caactggaaa tgtaaatatt tgctaatatt tctactatag gagaattaaa   4123

gtgagtgaat atggtaccac aaggtttgga gatttaattg ttgcaatgct gcatggatgg   4183

catatacacc aaacattcaa taattcttga ggataataat ggtaccacac aagatttgag   4243

gtgcatgaac gtcacgtgga caaaaggttt agtaattttt caagacaaca atgttaccac   4303

acacaagttt tgaggtgcat gcatggatgc cctgtggaaa gtttaaaaat attttggaaa   4363

tgatttgcat ggaagccatg tgtaaaacca tgacatccac ttggaggatg caataatgaa   4423

gaaaactaca aatttacatg caactagtta tgcatgtagt ctatataatg aggattttgc   4483

aatactttca ttcatacaca ctcactaagt tttacacgat tataatttct tcatagccag   4543

cagatctaaa atg ggc aaa gga ggg gac gct cgg gcc tcg aag ggc tca   4592
            Met Gly Lys Gly Gly Asp Ala Arg Ala Ser Lys Gly Ser
                575   580   585

acg gcg gct cgc aag atc agt tgg cag gaa gtc aag acc cac gcg tct   4640
Thr Ala Ala Arg Lys Ile Ser Trp Gln Glu Val Lys Thr His Ala Ser
            590   595   600

ccg gag gac gcc tgg atc att cac tcc aat aag gtc tac gac gtg tcc   4688
Pro Glu Asp Ala Trp Ile Ile His Ser Asn Lys Val Tyr Asp Val Ser
        605   610   615

aac tgg cac gaa cat ccc gga ggc gcc gtc att ttc acg cac gcc ggt   4736
Asn Trp His Glu His Pro Gly Gly Ala Val Ile Phe Thr His Ala Gly
        620   625   630

gac gac atg acg gac att ttc gct gcc ttt cac gca ccc gga tcg cag   4784
Asp Asp Met Thr Asp Ile Phe Ala Ala Phe His Ala Pro Gly Ser Gln
635   640   645

tcg ctc atg aag aag ttc tac att ggc gaa ttg ctc ccg gaa acc acc   4832
Ser Leu Met Lys Lys Phe Tyr Ile Gly Glu Leu Leu Pro Glu Thr Thr
650   655   660   665

ggc aag gag ccg cag caa atc gcc ttt gaa aag ggc tac cgc gat ctg   4880
Gly Lys Glu Pro Gln Gln Ile Ala Phe Glu Lys Gly Tyr Arg Asp Leu
            670   675   680

cgc tcc aaa ctc atc atg atg ggc atg ttc aag tcc aac aag tgg ttc   4928
Arg Ser Lys Leu Ile Met Met Gly Met Phe Lys Ser Asn Lys Trp Phe
            685   690   695

tac gtc tac aag tgc ctc agc aac atg gcc att tgg gcc gcc gcc tgt   4976
Tyr Val Tyr Lys Cys Leu Ser Asn Met Ala Ile Trp Ala Ala Ala Cys
        700   705   710

gct ctc gtc ttt tac tcg gac cgc ttc tgg gta cac ctg gcc agc gcc   5024

```
Ala Leu Val Phe Tyr Ser Asp Arg Phe Trp Val His Leu Ala Ser Ala
    715                 720             725
```

```
gtc atg ctg gga aca ttc ttt cag cag tcg gga tgg ttg gca cac gac        5072
Val Met Leu Gly Thr Phe Phe Gln Gln Ser Gly Trp Leu Ala His Asp
730             735             740             745
```

```
ttt ctg cac cac cag gtc ttc acc aag cgc aag cac ggg gat ctc gga        5120
Phe Leu His His Gln Val Phe Thr Lys Arg Lys His Gly Asp Leu Gly
                750             755             760
```

```
gga ctc ttt tgg ggg aac ctc atg cag ggt tac tcc gta cag tgg tgg        5168
Gly Leu Phe Trp Gly Asn Leu Met Gln Gly Tyr Ser Val Gln Trp Trp
            765             770             775
```

```
aaa aac aag cac aac gga cac cac gcc gtc ccc aac ctc cac tgc tcc        5216
Lys Asn Lys His Asn Gly His His Ala Val Pro Asn Leu His Cys Ser
        780             785             790
```

```
tcc gca gtc gcg caa gat ggg gac ccg gac atc gat acc atg ccc ctt        5264
Ser Ala Val Ala Gln Asp Gly Asp Pro Asp Ile Asp Thr Met Pro Leu
    795             800             805
```

```
ctc gcc tgg tcc gtc cag caa gcc cag tct tac cgg gaa ctc caa gcc        5312
Leu Ala Trp Ser Val Gln Gln Ala Gln Ser Tyr Arg Glu Leu Gln Ala
810             815             820             825
```

```
gac gga aag gat tcg ggt ttg gtc aag ttc atg atc cgt aac caa tcc        5360
Asp Gly Lys Asp Ser Gly Leu Val Lys Phe Met Ile Arg Asn Gln Ser
                830             835             840
```

```
tac ttt tac ttt ccc atc ttg ttg ctc gcc cgc ctg tcg tgg ttg aac        5408
Tyr Phe Tyr Phe Pro Ile Leu Leu Leu Ala Arg Leu Ser Trp Leu Asn
            845             850             855
```

```
gag tcc ttc aag tgc gcc ttt ggg ctt gga gct gcg tcg gag aac gct        5456
Glu Ser Phe Lys Cys Ala Phe Gly Leu Gly Ala Ala Ser Glu Asn Ala
        860             865             870
```

```
gct ctc gaa ctc aag gcc aag ggt ctt cag tac ccc ctt ttg gaa aag        5504
Ala Leu Glu Leu Lys Ala Lys Gly Leu Gln Tyr Pro Leu Leu Glu Lys
    875             880             885
```

```
gct ggc atc ctg ctg cac tac gct tgg atg ctt aca gtt tcg tcc ggc        5552
Ala Gly Ile Leu Leu His Tyr Ala Trp Met Leu Thr Val Ser Ser Gly
890             895             900             905
```

```
ttt gga cgc ttc tcg ttc gcg tac acc gca ttt tac ttt cta acc gcg        5600
Phe Gly Arg Phe Ser Phe Ala Tyr Thr Ala Phe Tyr Phe Leu Thr Ala
                910             915             920
```

```
acc gcg tcc tgt gga ttc ttg ctc gcc att gtc ttt ggc ctc ggc cac        5648
Thr Ala Ser Cys Gly Phe Leu Leu Ala Ile Val Phe Gly Leu Gly His
            925             930             935
```

```
aac ggc atg gcc acc tac aat gcc gac gcc cgt ccg gac ttc tgg aag        5696
Asn Gly Met Ala Thr Tyr Asn Ala Asp Ala Arg Pro Asp Phe Trp Lys
```

```
            940                    945                    950

ctc caa gtc acc acg act cgc aac gtc acg ggc gga cac ggt ttc ccc      5744
Leu Gln Val Thr Thr Thr Arg Asn Val Thr Gly Gly His Gly Phe Pro
    955                    960                    965

caa gcc ttt gtc gac tgg ttc tgt ggt ggc ctc cag tac caa gtc gac      5792
Gln Ala Phe Val Asp Trp Phe Cys Gly Gly Leu Gln Tyr Gln Val Asp
970                    975                    980                    985

cac cac tta ttc ccc agc ctg ccc cga cac aat ctg gcc aag aca  cac     5840
His His Leu Phe Pro Ser Leu Pro Arg His Asn Leu Ala Lys Thr   His
                   990                    995                    1000

gca ctg gtc gaa  tcg ttc tgc aag gag  tgg ggt gtc cag tac  cac       5885
Ala Leu Val Glu Ser Phe Cys Lys Glu Trp Gly Val Gln Tyr His
                   1005                   1010                   1015

gaa gcc gac ctt  gtg gac ggg acc atg  gaa gtc ttg cac cat  ttg       5930
Glu Ala Asp Leu Val Asp Gly Thr Met Glu Val Leu His His Leu
                   1020                   1025                   1030

ggc agc gtg gcc  ggc gaa ttc gtc gtg  gat ttt gta cgc gat  gga       5975
Gly Ser Val Ala Gly Glu Phe Val Val Asp Phe Val Arg Asp Gly
                   1035                   1040                   1045

ccc gcc atg taa agatctgccg gcatcgatcc cgggccatgg cctgctttaa          6027
Pro Ala Met
```

```
tgagatatgc gagacgccta tgatcgcatg atatttgctt tcaattctgt tgtgcacgtt    6087

gtaaaaaacc tgagcatgtg tagctcagat ccttaccgcc ggtttcggtt cattctaatg    6147

aatatatcac ccgttactat cgtattttta tgaataatat tctccgttca atttactgat    6207

tgtccgtcga cgagctcggc gcgccgtcga cctgcaggca tgcaagcttc acgctgccgc    6267

aagcactcag ggcgcaaggg ctgctaaagg aagcggaaca cgtagaaagc cagtccgcag    6327

aaacggtgct daccccggat gaatgtcagc tactgggcta tctggacaag ggaaaacgca    6387

agcgcaaaga gaaagcaggt agcttgcagt gggcttacat ggcgatagct agactgggcg    6447

gttttatgga cagcaagcga accggaattg ccagctgggg cgccctctgg taaggttggg    6507

aagccctgca aagtaaactg gatggctttc ttgccgccaa ggatctgatg gcgcagggga    6567

tcaagatcat gagcggagaa ttaagggagt cacgttatga cccccgccga tgacgcggga    6627

caagccgttt tacgtttgga actgacagaa ccgcaacgtt gaaggagcca ctcagccgcg    6687

ggtttctgga gtttaatgag ctaagcacat acgtcagaaa ccattattgc gcgttcaaaa    6747

gtcgcctaag gtcactatca gctagcaaat atttcttgtc aaaaatgctc cactgacgtt    6807

ccataaattc ccctcggtat ccaattagag tctcatattc actctcaatc cagatctcga    6867
```

```
ctctagtcga gggcccatgg gagcttggat tgaacaagat ggattgcacg caggttctcc    6927

ggccgcttgg gtggagaggc tattcggcta tgactgggca caacagacaa tcggctgctc    6987

tgatgccgcc gtgttccggc tgtcagcgca ggggcgcccg gttctttttg tcaagaccga    7047

cctgtccggt gccctgaatg aactgcagga cgaggcagcg cggctatcgt ggctggccac    7107

gacgggcgtt ccttgcgcag ctgtgctcga cgttgtcact gaagcgggaa gggactggct    7167

gctattgggc gaagtgccgg ggcaggatct cctgtcatct caccttgctc ctgccgagaa    7227

agtatccatc atggctgatg caatgcggcg gctgcatacg cttgatccgg ctacctgccc    7287

attcgaccac caagcgaaac atcgcatcga gcgagcacgt actcggatgg aagccggtct    7347

tgtcgatcag gatgatctgg acgaagagca tcaggggctc gcgccagccg aactgttcgc    7407

caggctcaag gcgcgcatgc ccgacggcga ggatctcgtc gtgacccatg gcgatgcctg    7467

cttgccgaat atcatggtgg aaaatggccg cttttctgga ttcatcgact gtggccggct    7527

gggtgtggcg gaccgctatc aggacatagc gttggctacc cgtgatattg ctgaagagct    7587

tggcggcgaa tgggctgacc gcttcctcgt gctttacggt atcgccgctc ccgattcgca    7647

gcgcatcgcc ttctatcgcc ttcttgacga gttcttctga gcgggaccca agctagcttc    7707

gacggatccc ccgatgagct aagctagcta tatcatcaat ttatgtatta cacataatat    7767

cgcactcagt ctttcatcta cggcaatgta ccagctgata taatcagtta ttgaaatatt    7827

tctgaattta aacttgcatc aataaattta tgtttttgct tggactataa tacctgactt    7887

gttattttat caataaatat ttaaactata tttctttcaa gatgggaatt aattcactgg    7947

ccgtcgtttt acaacgtcgt gactgggaaa accctggcgt tacccaactt aatcgccttg    8007

cagcacatcc ccctttcgcc agctggcgta atagcgaaga ggcccgcacc gatcgccctt    8067

cccaacagtt gcgcagcctg aatggcgccc gctcctttcg ctttcttccc ttcctttctc    8127

gccacgttcg ccggctttcc ccgtcaagct ctaaatcggg ggctcccttt agggttccga    8187

tttagtgctt tacggcacct cgaccccaaa aaacttgatt tgggtgatgg ttcacgtagt    8247

gggccatcgc cctgatagac ggtttttcgc cctttgacgt tggagtccac gttctttaat    8307

agtggactct tgttccaaac tggaacaaca ctcaacccta tctcgggcta ttcttttgat    8367

ttataaggga ttttgccgat ttcggaacca ccatcaaaca ggattttcgc ctgctggggc    8427

aaaccagcgt ggaccgcttg ctgcaactct ctcagggcca ggcggtgaag gcaatcagc    8487

tgttgcccgt ctcactggtg aaaagaaaaa ccaccccagt acattaaaaa cgtccgcaat    8547
```

```
gtgttattaa gttgtctaag cgtcaatttg tttacaccac aatatatcct gccaccagcc    8607

agccaacagc tccccgaccg gcagctcggc acaaaatcac cactcgatac aggcagccca    8667

tcagtccggg acggcgtcag cgggagagcc gttgtaaggc ggcagacttt gctcatgtta    8727

ccgatgctat tcggaagaac ggcaactaag ctgccgggtt tgaaacacgg atgatctcgc    8787

ggagggtagc atgttgattg taacgatgac agagcgttgc tgcctgtgat caaatatcat    8847

ctccctcgca gagatccgaa ttatcagcct tcttattcat ttctcgctta accgtgacag    8907

gctgtcgatc ttgagaacta tgccgacata ataggaaatc gctggataaa gccgctgagg    8967

aagctgagtg gcgctatttc tttagaagtg aacgttgacg atatcaactc ccctatccat    9027

tgctcaccga atggtacagg tcggggaccc gaagttccga ctgtcggcct gatgcatccc    9087

cggctgatcg accccagatc tggggctgag aaagcccagt aaggaaacaa ctgtaggttc    9147

gagtcgcgag atcccccgga accaaaggaa gtaggttaaa cccgctccga tcaggccgag    9207

ccacgccagg ccgagaacat tggttcctgt aggcatcggg attggcggat caaacactaa    9267

agctactgga acgagcagaa gtcctccggc cgccagttgc caggcggtaa aggtgagcag    9327

aggcacggga ggttgccact tgcgggtcag cacggttccg aacgccatgg aaaccgcccc    9387

cgccaggccc gctgcgacgc cgacaggatc tagcgctgcg tttggtgtca acaccaacag    9447

cgccacgccc gcagttccgc aaatagcccc caggaccgcc atcaatcgta tcgggctacc    9507

tagcagagcg gcagagatga acacgaccat cagcggctgc acagcgccta ccgtcgccgc    9567

gaccccgccc ggcaggcggt agaccgaaat aaacaacaag ctccagaata gcgaaatatt    9627

aagtgcgccg aggatgaaga tgcgcatcca ccagattccc gttggaatct gtcggacgat    9687

catcacgagc aataaacccg ccggcaacgc ccgcagcagc ataccggcga cccctcggcc    9747

tcgctgttcg ggctccacga aaacgccgga cagatgcgcc ttgtgagcgt ccttggggcc    9807

gtcctcctgt ttgaagaccg acagcccaat gatctcgccg tcgatgtagg cgccgaatgc    9867

cacggcatct cgcaaccgtt cagcgaacgc ctccatgggc tttttctcct cgtgctcgta    9927

aacggacccg aacatctctg gagctttctt cagggccgac aatcggatct cgcggaaatc    9987

ctgcacgtcg gccgctccaa gccgtcgaat ctgagcctta atcacaattg tcaattttaa   10047

tcctctgttt atcggcagtt cgtagagcgc gccgtgcgtc ccgagcgata ctgagcgaag   10107

caagtgcgtc gagcagtgcc cgcttgttcc tgaaatgcca gtaaagcgct ggctgctgaa   10167

cccccagccg gaactgaccc cacaaggccc tagcgtttgc aatgcaccag gtcatcattg   10227

acccaggcgt gttccaccag gccgctgcct cgcaactctt cgcaggcttc gccgacctgc   10287
```

```
tcgcgccact tcttcacgcg ggtggaatcc gatccgcaca tgaggcggaa ggtttccagc   10347

ttgagcgggt acggctcccg gtgcgagctg aaatagtcga acatccgtcg ggccgtcggc   10407

gacagcttgc ggtacttctc ccatatgaat ttcgtgtagt ggtcgccagc aaacagcacg   10467

acgatttcct cgtcgatcag gacctggcaa cgggacgttt tcttgccacg gtccaggacg   10527

cggaagcggt gcagcagcga caccgattcc aggtgcccaa cgcggtcgga cgtgaagccc   10587

atcgccgtcg cctgtaggcg cgacaggcat tcctcggcct tcgtgtaata ccggccattg   10647

atcgaccagc ccaggtcctg gcaaagctcg tagaacgtga aggtgatcgg ctcgccgata   10707

ggggtgcgct tcgcgtactc caacacctgc tgccacacca gttcgtcatc gtcggcccgc   10767

agctcgacgc cggtgtaggt gatcttcacg tccttgttga cgtggaaaat gaccttgttt   10827

tgcagcgcct cgcgcgggat tttcttgttg cgcgtggtga acagggcaga gcgggccgtg   10887

tcgtttggca tcgctcgcat cgtgtccggc cacggcgcaa tatcgaacaa ggaaagctgc   10947

atttccttga tctgctgctt cgtgtgtttc agcaacgcgg cctgcttggc ctcgctgacc   11007

tgttttgcca ggtcctcgcc ggcggttttt cgcttcttgg tcgtcatagt tcctcgcgtg   11067

tcgatggtca tcgacttcgc caaacctgcc gcctcctgtt cgagacgacg cgaacgctcc   11127

acggcggccg atggcgcggg cagggcaggg ggagccagtt gcacgctgtc gcgctcgatc   11187

ttggccgtag cttgctggac catcgagccg acggactgga aggtttcgcg gggcgcacgc   11247

atgacggtgc ggcttgcgat ggtttcggca tcctcggcgg aaaaccccgc gtcgatcagt   11307

tcttgcctgt atgccttccg gtcaaacgtc cgattcattc accctccttg cgggattgcc   11367

ccgactcacg ccggggcaat gtgcccttat tcctgatttg acccgcctgg tgccttggtg   11427

tccagataat ccaccttatc ggcaatgaag tcggtcccgt agaccgtctg gccgtccttc   11487

tcgtacttgg tattccgaat cttgccctgc acgaatacca gcgacccctt gcccaaatac   11547

ttgccgtggg cctcggcctg agagccaaaa cacttgatgc ggaagaagtc ggtgcgctcc   11607

tgcttgtcgc cggcatcgtt gcgccacatc taggtactaa aacaattcat ccagtaaaat   11667

ataatatttt attttctccc aatcaggctt gatccccagt aagtcaaaaa atagctcgac   11727

atactgttct tccccgatat cctccctgat cgaccggacg cagaaggcaa tgtcatacca   11787

cttgtccgcc ctgccgcttc tcccaagatc aataaagcca cttactttgc catctttcac   11847

aaagatgttg ctgtctccca ggtcgccgtg ggaaaagaca agttcctctt cgggcttttc   11907

cgtctttaaa aaatcataca gctcgcgcgg atctttaaat ggagtgtctt cttcccagtt   11967
```

```
ttcgcaatcc acatcggcca gatcgttatt cagtaagtaa tccaattcgg ctaagcggct   12027

gtctaagcta ttcgtatagg gacaatccga tatgtcgatg gagtgaaaga gcctgatgca   12087

ctccgcatac agctcgataa tcttttcagg gctttgttca tcttcatact cttccgagca   12147

aaggacgcca tcggcctcac tcatgagcag attgctccag ccatcatgcc gttcaaagtg   12207

caggaccttt ggaacaggca gctttccttc cagccatagc atcatgtcct tttcccgttc   12267

cacatcatag gtggtccctt tataccggct gtccgtcatt tttaaatata ggttttcatt   12327

ttctcccacc agcttatata ccttagcagg agacattcct tccgtatctt ttacgcagcg   12387

gtatttttcg atcagttttt tcaattccgg tgatattctc attttagcca tttattattt   12447

ccttcctctt ttctacagta tttaaagata ccccaagaag ctaattataa caagacgaac   12507

tccaattcac tgttccttgc attctaaaac cttaaatacc agaaaacagc tttttcaaag   12567

ttgttttcaa agttggcgta taacatagta tcgacggagc cgattttgaa accacaatta   12627

tgggtgatgc tgccaactta ctgatttagt gtatgatggt gtttttgagg tgctccagtg   12687

gcttctgtgt ctatcagctg tccctcctgt tcagctactg acggggtggt gcgtaacggc   12747

aaaagcaccg ccggacatca gcgctatctc tgctctcact gccgtaaaac atggcaactg   12807

cagttcactt acaccgcttc tcaacccggt acgcaccaga aaatcattga tatggccatg   12867

aatggcgttg gatgccgggc aacagcccgc attatgggcg ttggcctcaa cacgatttta   12927

cgtcacttaa aaaactcagg ccgcagtcgg taacctcgcg catacagccg ggcagtgacg   12987

tcatcgtctg cgcggaaatg gacgaacagt ggggctatgt cggggctaaa tcgcgccagc   13047

gctggctgtt ttacgcgtat gacagtctcc ggaagacggt tgttgcgcac gtattcggtg   13107

aacgcactat ggcgacgctg gggcgtctta tgagcctgct gtcacccttt gacgtggtga   13167

tatggatgac ggatggctgg ccgctgtatg aatcccgcct gaagggaaag ctgcacgtaa   13227

tcagcaagcg atatacgcag cgaattgagc ggcataacct gaatctgagg cagcacctgg   13287

cacggctggg acggaagtcg ctgtcgttct caaaatcggt ggagctgcat gacaaagtca   13347

tcgggcatta tctgaacata aaacactatc aataagttgg agtcattacc caattatgat   13407

agaatttaca agctataagg ttattgtcct gggtttcaag cattagtcca tgcaagtttt   13467

tatgctttgc ccattctata gatatattga taagcgcgct gcctatgcct tgccccctga   13527

aatccttaca tacggcgata tcttctatat aaaagatata ttatcttatc agtattgtca   13587

atatattcaa ggcaatctgc ctcctcatcc tcttcatcct cttcgtcttg gtagcttttt   13647

aaatatggcg cttcatagag taattctgta aaggtccaat tctcgttttc atacctcggt   13707
```

```
ataatcttac ctatcacctc aaatggttcg ctgggtttat cgcacccccg aacacgagca    13767

cggcacccgc gaccactatg ccaagaatgc ccaaggtaaa aattgccggc cccgccatga    13827

agtccgtgaa tgccccgacg gccgaagtga agggcaggcc gccacccagg ccgccgccct    13887

cactgcccgg cacctggtcg ctgaatgtcg atgccagcac ctgcggcacg tcaatgcttc    13947

cgggcgtcgc gctcgggctg atcgcccatc ccgttactgc cccgatcccg gcaatggcaa    14007

ggactgccag cgctgccatt tttggggtga ggccgttcgc ggccgagggg cgcagcccct    14067

gggggggatgg gaggcccgcg ttagcgggcc gggagggttc gagaaggggg ggcacccccc    14127

ttcggcgtgc gcggtcacgc gcacagggcg cagccctggt taaaaacaag gtttataaat    14187

attggtttaa aagcaggtta aaagacaggt tagcggtggc cgaaaaacgg gcggaaaccc    14247

ttgcaaatgc tggattttct gcctgtggac agcccctcaa atgtcaatag gtgcgcccct    14307

catctgtcag cactctgccc ctcaagtgtc aaggatcgcg cccctcatct gtcagtagtc    14367

gcgcccctca agtgtcaata ccgcagggca cttatcccca ggcttgtcca catcatctgt    14427

gggaaactcg cgtaaaatca ggcgttttcg ccgatttgcg aggctggcca gctccacgtc    14487

gccggccgaa atcgagcctg cccctcatct gtcaacgccg cgccgggtga gtcggcccct    14547

caagtgtcaa cgtccgcccc tcatctgtca gtgagggcca agttttccgc gaggtatcca    14607

caacgccggc ggccgcggtg tctcgcacac ggcttcgacg gcgtttctgg cgcgtttgca    14667

gggccataga cggccgccag cccagcggcg agggcaacca gcccggtgag cgtcgcaaag    14727

gcgctcggtc ttgccttgct cgtcggtgat gtacttcacc agctccgcga agtcgctctt    14787

cttgatggag cgcatgggga cgtgcttggc aatcacgcgc accccccggc cgttttagcg    14847

gctaaaaaag tcatggctct gccctcgggc ggaccacgcc catcatgacc ttgccaagct    14907

cgtcctgctt ctcttcgatc ttcgccagca gggcgaggat cgtggcatca ccgaaccgcg    14967

ccgtgcgcgg gtcgtcggtg agccagagtt tcagcaggcc gcccaggcgg cccaggtcgc    15027

cattgatgcg gccagctcg cggacgtgct catagtccac gacgcccgtg attttgtagc    15087

cctggccgac ggccagcagg taggccgaca ggctcatgcc ggccgccgcc gcctttttcct    15147

caatcgctct tcgttcgtct ggaaggcagt acaccttgat aggtgggctg cccttcctgg    15207

ttggcttggt ttcatcagcc atccgcttgc cctcatctgt tacgccggcg gtagccggcc    15267

agcctcgcag agcaggattc ccgttgagca ccgccaggtg cgaataaggg acagtgaaga    15327

aggaacaccc gctcgcgggt gggcctactt cacctatcct gcccggctga cgccgttgga    15387
```

```
tacaccaagg aaagtctaca cgaacccttt ggcaaaatcc tgtatatcgt gcgaaaaagg   15447

atggatatac cgaaaaaatc gctataatga ccccgaagca gggttatgca gcggaaaagc   15507

gccacgcttc ccgaagggag aaaggcggac aggtatccgg taagcggcag ggtcggaaca   15567

ggagagcgca cgagggagct tccaggggga aacgcctggt atctttatag tcctgtcggg   15627

tttcgccacc tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg gcggagccta   15687

tggaaaaacg ccagcaacgc ggccttttta cggttcctgg ccttttgctg gccttttgct   15747

cacatgttct ttcctgcgtt atcccctgat tctgtggata accgtattac cgcctttgag   15807

tgagctgata ccgctcgccg cagccgaacg accgagcgca gcgagtcagt gagcgaggaa   15867

gcggaagagc gccagaaggc cgccagagag gccgagcgcg gccgtgaggc ttggacgcta   15927

gggcagggca tgaaaaagcc cgtagcgggc tgctacgggc gtctgacgcg gtggaaaggg   15987

ggaggggatg ttgtctacat ggctctgctg tagtgagtgg gttgcgctcc ggcagcggtc   16047

ctgatcaatc gtcacccttt ctcggtcctt caacgttcct gacaacgagc ctccttttcg   16107

ccaatccatc gacaatcacc gcgagtccct gctcgaacgc tgcgtccgga ccggcttcgt   16167

cgaaggcgtc tatcgcggcc cgcaacagcg gcgagagcgg agcctgttca acggtgccgc   16227

cgcgctcgcc ggcatcgctg tcgccggcct gctcctcaag cacggcccca acagtgaagt   16287

agctgattgt catcagcgca ttgacggcgt ccccggccga aaaacccgcc tcgcagagga   16347

agcgaagctg cgcgtcggcc gtttccatct gcggtgcgcc cggtcgcgtg ccggcatgga   16407

tgcgcgcgcc atcgcggtag gcgagcagcg cctgcctgaa gctgcgggca ttcccgatca   16467

gaaatgagcg ccagtcgtcg tcggctctcg gcaccgaatg cgtatgattc tccgccagca   16527

tggcttcggc cagtgcgtcg agcagcgccc gcttgttcct gaagtgccag taaagcgccg   16587

gctgctgaac ccccaaccgt tccgccagtt tgcgtgtcgt cagaccgtct acgccgacct   16647

cgttcaacag gtccagggcg gcacggatca ctgtattcgg ctgcaacttt gtcatgcttg   16707

acactttatc actgataaac ataatatgtc caccaactta tcagtgataa agaatccgcg   16767

cgttcaatcg gaccagcgga ggctggtccg gaggccagac gtgaaaccca acatacccct   16827

gatcgtaatt ctgagcactg tcgcgctcga cgctgtcggc atcggcctga ttatgccggt   16887

gctgccgggc ctcctgcgcg atctggttca ctcgaacgac gtcaccgccc actatggcat   16947

tctgctggcg ctgtatgcgt tggtgcaatt tgcctgcgca cctgtgctgg gcgcgctgtc   17007

ggatcgtttc gggcggcggc caatcttgct cgtctcgctg gccggcgcca gatc          17061
```

<210> 35
<211> 290
<212> PRT
<213> Phaeodactylum tricornutum, Physcomitrella patens, Caenorhabditis elegans

<400> 35

Met Glu Val Val Glu Arg Phe Tyr Gly Glu Leu Asp Gly Lys Val Ser
1   5   10   15

Gln Gly Val Asn Ala Leu Leu Gly Ser Phe Gly Val Glu Leu Thr Asp
            20   25   30

Thr Pro Thr Thr Lys Gly Leu Pro Leu Val Asp Ser Pro Thr Pro Ile
        35   40   45

Val Leu Gly Val Ser Val Tyr Leu Thr Ile Val Ile Gly Gly Leu Leu
        50   55   60

Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg Ala Ser Glu Pro Phe Leu
65   70   75   80

Leu Gln Ala Leu Val Leu Val His Asn Leu Phe Cys Phe Ala Leu Ser
            85   90   95

Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln Ala Ile Thr Trp Arg Tyr
            100   105   110

Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys His Lys Glu Met Ala Ile
        115   120   125

Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr Val Glu Phe Met Asp Thr
        130   135   140

Val Ile Met Ile Leu Lys Arg Ser Thr Arg Gln Ile Ser Phe Leu His
145   150   155   160

Val Tyr His His Ser Ser Ile Ser Leu Ile Trp Trp Ala Ile Ala His
            165   170   175

His Ala Pro Gly Gly Glu Ala Tyr Trp Ser Ala Ala Leu Asn Ser Gly
            180   185   190

Val His Val Leu Met Tyr Ala Tyr Tyr Phe Leu Ala Ala Cys Leu Arg
            195   200   205

Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu Phe Trp Gly Arg Tyr Leu
    210   215   220

Thr Gln Phe Gln Met Phe Gln Phe Met Leu Asn Leu Val Gln Ala Tyr
225   230   235   240

Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro Gln Trp Leu Ile Lys Ile
              245   250   255

Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe Leu Phe Gly Asn Phe Tyr
          260   265   270

Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly Lys Gln Lys Gly Ala Lys
      275   280   285

Thr Glu
    290

<210> 36
<211> 282
<212> PRT
<213> Phaeodactylum tricornutum, Physcomitrella patens, Caenorhabditis elegans

<400> 36

Met Glu Asn Phe Trp Ser Ile Val Val Phe Phe Leu Leu Ser Ile Leu
1   5   10   15

Phe Ile Leu Tyr Asn Ile Ser Thr Val Cys His Tyr Tyr Met Arg Ile
          20   25   30

Ser Phe Tyr Tyr Phe Thr Ile Leu Leu His Gly Met Glu Val Cys Val
        35   40   45

Thr Met Ile Pro Ser Trp Leu Asn Gly Lys Gly Ala Asp Tyr Val Phe
    50   55   60

His Ser Phe Phe Tyr Trp Cys Lys Trp Thr Gly Val His Thr Thr Val
65   70   75   80

Tyr Gly Tyr Glu Lys Thr Gln Val Glu Gly Pro Ala Val Val Ile Cys
              85   90   95

Asn His Gln Ser Ser Leu Asp Ile Leu Ser Met Ala Ser Ile Trp Pro
           100   105   110

Lys Asn Cys Val Val Met Met Lys Arg Ile Leu Ala Tyr Val Pro Phe
          115   120   125

Phe Asn Leu Gly Ala Tyr Phe Ser Asn Thr Ile Phe Ile Asp Arg Tyr
   130   135   140

Asn Arg Glu Arg Ala Met Ala Ser Val Asp Tyr Cys Ala Ser Glu Met
145   150   155   160

Lys Asn Arg Asn Leu Lys Leu Trp Val Phe Pro Glu Gly Thr Arg Asn
          165   170   175

Arg Glu Gly Gly Phe Ile Pro Phe Lys Lys Gly Ala Phe Asn Ile Ala
          180   185   190

Val Arg Ala Gln Ile Pro Ile Ile Pro Val Val Phe Ser Asp Tyr Arg
          195   200   205

Asp Phe Tyr Ser Lys Pro Gly Arg Tyr Phe Lys Asn Asp Gly Glu Val
   210   215   220

Val Ile Arg Val Leu Asp Ala Ile Pro Thr Lys Gly Leu Thr Leu Asp
225   230   235   240

Asp Val Ser Glu Leu Ser Asp Met Cys Arg Asp Val Met Leu Ala Ala
          245   250   255

Tyr Lys Glu Val Thr Leu Glu Ala Gln Gln Arg Asn Ala Thr Arg Arg
          260   265   270

Gly Glu Thr Lys Asp Gly Lys Lys Ser Glu
          275   280

<210> 37
<211> 477
<212> PRT
<213> Phaeodactylum tricornutum, Physcomitrella patens, Caenorhabditis elegans

<400> 37

Met Gly Lys Gly Gly Asp Ala Arg Ala Ser Lys Gly Ser Thr Ala Ala

```
 1                    5                    10                   15

Arg Lys Ile Ser Trp Gln Glu Val Lys Thr His Ala Ser Pro Glu Asp
            20                  25                  30

Ala Trp Ile Ile His Ser Asn Lys Val Tyr Asp Val Ser Asn Trp His
        35                  40                  45

Glu His Pro Gly Gly Ala Val Ile Phe Thr His Ala Gly Asp Asp Met
    50                  55                  60

Thr Asp Ile Phe Ala Ala Phe His Ala Pro Gly Ser Gln Ser Leu Met
65                      70                  75                  80

Lys Lys Phe Tyr Ile Gly Glu Leu Leu Pro Glu Thr Thr Gly Lys Glu
            85                  90                  95

Pro Gln Gln Ile Ala Phe Glu Lys Gly Tyr Arg Asp Leu Arg Ser Lys
            100                 105                 110

Leu Ile Met Met Gly Met Phe Lys Ser Asn Lys Trp Phe Tyr Val Tyr
            115                 120                 125

Lys Cys Leu Ser Asn Met Ala Ile Trp Ala Ala Ala Cys Ala Leu Val
    130                 135                 140

Phe Tyr Ser Asp Arg Phe Trp Val His Leu Ala Ser Ala Val Met Leu
145                 150                 155                 160

Gly Thr Phe Phe Gln Gln Ser Gly Trp Leu Ala His Asp Phe Leu His
            165                 170                 175

His Gln Val Phe Thr Lys Arg Lys His Gly Asp Leu Gly Gly Leu Phe
            180                 185                 190

Trp Gly Asn Leu Met Gln Gly Tyr Ser Val Gln Trp Trp Lys Asn Lys
        195                 200                 205

His Asn Gly His His Ala Val Pro Asn Leu His Cys Ser Ser Ala Val
    210                 215                 220

Ala Gln Asp Gly Asp Pro Asp Ile Asp Thr Met Pro Leu Leu Ala Trp
```

225   230   235   240

Ser Val Gln Gln Ala Gln Ser Tyr Arg Glu Leu Gln Ala Asp Gly Lys
            245   250   255

Asp Ser Gly Leu Val Lys Phe Met Ile Arg Asn Gln Ser Tyr Phe Tyr
            260   265   270

Phe Pro Ile Leu Leu Leu Ala Arg Leu Ser Trp Leu Asn Glu Ser Phe
            275   280   285

Lys Cys Ala Phe Gly Leu Gly Ala Ala Ser Glu Asn Ala Ala Leu Glu
            290   295   300

Leu Lys Ala Lys Gly Leu Gln Tyr Pro Leu Leu Glu Lys Ala Gly Ile
305   310   315   320

Leu Leu His Tyr Ala Trp Met Leu Thr Val Ser Ser Gly Phe Gly Arg
            325   330   335

Phe Ser Phe Ala Tyr Thr Ala Phe Tyr Phe Leu Thr Ala Thr Ala Ser
            340   345   350

Cys Gly Phe Leu Leu Ala Ile Val Phe Gly Leu Gly His Asn Gly Met
            355   360   365

Ala Thr Tyr Asn Ala Asp Ala Arg Pro Asp Phe Trp Lys Leu Gln Val
            370   375   380

Thr Thr Thr Arg Asn Val Thr Gly Gly His Gly Phe Pro Gln Ala Phe
385   390   395   400

Val Asp Trp Phe Cys Gly Gly Leu Gln Tyr Gln Val Asp His His Leu
                405   410   415

Phe Pro Ser Leu Pro Arg His Asn Leu Ala Lys Thr His Ala Leu Val
            420   425   430

Glu Ser Phe Cys Lys Glu Trp Gly Val Gln Tyr His Glu Ala Asp Leu
            435   440   445

Val Asp Gly Thr Met Glu Val Leu His His Leu Gly Ser Val Ala Gly
            450   455   460

```
Glu Phe Val Val Asp Phe Val Arg Asp Gly Pro Ala Met
465                 470                 475
```

```
<210>  38
<211>  47
<212>  DNA
<213>  artificial sequence

<220>
<221>  misc_feature
<222>  (1)..(47)
<223>  synthetic oligonucleotide

<400>  38
ccggaattcg gcgcgccgag ctcctcgagc aaatttacac attgcca                47
```

```
<210>  39
<211>  47
<212>  DNA
<213>  artificial sequence

<220>
<221>  misc_feature
<222>  (1)..(47)
<223>  synthetic oligonucleotide

<400>  39
ccggaattcg gcgcgccgag ctcctcgagc aaatttacac attgcca                47
```

```
<210>  40
<211>  47
<212>  DNA
<213>  artificial sequence

<220>
<221>  misc_feature
<222>  (1)..(47)
<223>  synthetic oligonucleotide

<400>  40
ccggaattcg gcgcgccgag ctcctcgagc aaatttacac attgcca                47
```

```
<210>  41
<211>  48
<212>  DNA
<213>  artificial sequence

<220>
<221>  misc_feature
<222>  (1)..(48)
```

<223>   synthetic oligonucleotide

<400>   41
aaaactgcag gcggccgccc accgcggtgg gctggctatg aagaaatt   48

<210>   42
<211>   27
<212>   DNA
<213>   artificial sequence

<220>
<221>   misc_feature
<222>   (1)..(27)
<223>   synthetic oligonucleotide

<400>   42
cgcggatccg ctggctatga agaaatt   27

<210>   43
<211>   45
<212>   DNA
<213>   artificial sequence

<220>
<221>   misc_feature
<222>   (1)..(45)
<223>   synthetic oligonucleotide

<400>   43
tcccccggga tcgatgccgg cagatctgct ggctatgaag aaatt   45

<210>   44
<211>   40
<212>   DNA
<213>   artificial sequence

<220>
<221>   misc_feature
<222>   (1)..(40)
<223>   synthetic oligonucleotide

<400>   44
aaaactgcag tctagaaggc ctcctgcttt aatgagatat   40

<210>   45
<211>   51
<212>   DNA
<213>   artificial sequence

<220>
<221>   misc_feature
<222>   (1)..(51)

<223> synthetic oligonucleotide

<400> 45
cgcggatccg atatcgggcc cgctagcgtt aaccctgctt taatgagata t   51

<210> 46
<211> 33
<212> DNA
<213> artificial sequence

<220>
<221> misc_feature
<222> (1)..(33)
<223> synthetic oligonucleotide

<400> 46
tcccccgggc catggcctgc tttaatgaga tat   33

<210> 47
<211> 53
<212> DNA
<213> artificial sequence

<220>
<221> misc_feature
<222> (1)..(53)
<223> synthetic oligonucleotide

<400> 47
cccaagcttg gcgcgccgag ctcgaattcg tcgacggaca atcagtaaat tga   53

<210> 48
<211> 53
<212> DNA
<213> artificial sequence

<220>
<221> misc_feature
<222> (1)..(53)
<223> synthetic oligonucleotide

<400> 48
cccaagcttg gcgcgccgag ctcgaattcg tcgacggaca atcagtaaat tga   53

<210> 49
<211> 47
<212> DNA
<213> artificial sequence

<220>
<221> misc_feature
<222> (1)..(47)
<223> synthetic oligonucleotide

<400> 49
cccaagcttg gcgcgccgag ctcgtcgacg gacaatcagt aaattga   47

<210> 50
<211> 29
<212> DNA
<213> artificial sequence

<220>
<221> misc_feature
<222> (1)..(29)
<223> synthetic oligonucleotide

<400> 50
acataatgga gaacttctgg tcgatcgtc   29

<210> 51
<211> 24
<212> DNA
<213> artificial sequence

<220>
<221> misc_feature
<222> (1)..(24)
<223> synthetic oligonucleotide

<400> 51
ttactcagat ttcttcccgt cttt   24

<210> 52
<211> 26
<212> DNA
<213> artificial sequence

<220>
<221> misc_feature
<222> (1)..(26)
<223> synthetic oligonucleotide

<400> 52
acataatgac cttcctagcc atatta   26

<210> 53
<211> 24
<212> DNA
<213> artificial sequence

<220>
<221> misc_feature
<222> (1)..(24)
<223> synthetic oligonucleotide

```
<400>  53
tcagatattc aaattggcgg cttc                                          24


<210>  54
<211>  32
<212>  DNA
<213>  artificial sequence

<220>
<221>  misc_feature
<222>  (1)..(32)
<223>  synthetic oligonucleotide

<400>  54
ttaagcgcgg ccgcatggag aacttctggt cg                                 32


<210>  55
<211>  31
<212>  DNA
<213>  artificial sequence

<220>
<221>  misc_feature
<222>  (1)..(31)
<223>  synthetic oligonucleotide

<400>  55
acctcggcgg ccgccctttt actcagattt c                                  31


<210>  56
<211>  41
<212>  DNA
<213>  artificial sequence

<220>
<221>  misc_feature
<222>  (1)..(41)
<223>  synthetic oligonucleotide

<400>  56
acataatgga gaacttctgg tctattgttg tgttttttct a                       41


<210>  57
<211>  41
<212>  DNA
<213>  artificial sequence

<220>
<221>  misc_feature
<222>  (1)..(41)
<223>  synthetic oligonucleotide

<400>  57
```

ctagctagct tactcagatt tcttcccgtc ttttgtttct c   41